Molecular Biology and Evolution, 2026, 43, 1–22
https://doi.org/10.1093/molbev/msag027
Published: 2 February 2026
Discoveries

MBE OXFORD

Dynamic regulation of spermatogenesis and hybrid sterility revealed by single-cell analysis in yak and cattle

Shi-Xin Wu [1,2†], Guo-Wen Wang [1,2†], Rui-Dong Wan [1,2†], Rong-Ge Yan [1,2], Xue Gao [1,2], Rui-Na Zhang [1], Qi-Lin Yang [3], You-Gui Fang [4], Gong-Xue Jia [1,2], Qi-En Yang [1,2,5*]

[1]Key Laboratory of Adaptation and Evolution of Plateau Biota, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining 810000, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China
[3]Department of Veterinary Sciences, Qinghai Vocational Technical College of Animal Science and Agriculture, Xining 810016, China
[4]Maduo Animal Science Service Center, Qinghai Province 813500, China
[5]Qinghai Key Laboratory of Animal Ecological Genomics, Northwest Institute of Plateau Biology, Chinese Academy of Sciences, Xining 810001, China
[†]These authors contributed equally to this work.
*Corresponding author: E-mail: yangqien@nwipb.cas.cn.
Associate editor: Aya Takahashi

Abstract

Spermatogenesis is a highly orchestrated germ cell differentiation process involving the dynamic regulation of cell fate transitions. Dissecting the molecular landscapes of spermatogenic cell types is crucial for identifying fertility-related problems and improving the reproductive performance of farm animals. Here, we conducted transcriptomic and chromosome spreading across meiotic stages of testicular cells from taurine cattle (*Bos taurus*), yak (*Bos grunniens*), and their hybrid progenies to describe the transcriptional landscape of normal spermatogenesis and identify potential regulators that are involved in hybrid sterility. The results revealed seven types of spermatogonia, ten spermatocytes and ten types of spermatids in the cattle or yak testes. In sharp contrast, the testes of the cattle–yak hybrids contained only seven spermatogonial subtypes and six types of spermatocytes. Notably, the arrest of spermatocytes at the diplotene-to-diakinesis transition was accompanied by defects in double-strand break repair. In the testes of backcrossed offspring, spermatogenic arrest was partially rescued, and round spermatozoa were produced. By performing joint analysis, we identified 115 genes that exhibited differential protein abundance in spermatocytes of cattle-yak. Among them, 24 genes carrying genomic structural variations were differentially expressed in spermatocytes of cattle-yak but recovered in those of backcrossed offspring. This work provides important insights into spermatogenesis in large animals and serves as a valuable resource for identifying the factors determining reproductive isolation.

Keywords spermatogenesis, cattle-yak, single-cell RNA-seq, meiosis, hybrid sterility, structural variation

Introduction

Male fertility relies on the continuous and robust production of functional sperm that transmit genetic and epigenetic information to the next generation. Spermatogenesis is a hierarchical cellular differentiation process involving mitosis of spermatogonia, meiosis of spermatocytes, and an elaborately organized process termed spermiogenesis (de Rooij 2017; Griswold 2018; Sohni et al. 2019). These elaborate organized events must be precisely regulated to ensure that the spermatogenic lineage is properly preserved and that large numbers of sperm are produced (Eddy 1998). The gene expression patterns and critical regulators involved in spermatogenesis have been extensively studied in mice, humans, and nonhuman primates via single-cell RNA sequencing (scRNA-seq) (Green et al. 2018; Shami et al. 2020; Di Persio et al. 2021). A recent study analyzed single-nucleus transcriptome data of testes from 11 species and identified conserved programs governing spermatogenesis during evolution (Murat et al. 2023). Livestock species offer unique models for understanding conserved and divergent mechanisms directing mammalian spermatogenesis. In addition, the fertility of cattle, which are major contributors to milk and beef consumption worldwide, has declined in the last two decades (Royal et al. 2000; Charlier et al. 2016). However, changes in cell composition and the dynamics of gene expression during spermatogenesis are largely unknown in these species.

Spermatogenesis in bulls involves unique patterns of gene expression, spermatogenic cell fate decisions, and germline-niche interactions within the seminiferous tubules (Staub and Johnson 2018; Gewaily et al. 2020). The specification of bovine primordial germ cells occurs 2 to 8 weeks after fetal development, and this event appears to be governed by conserved

Received: June 5, 2025. **Revised:** November 30, 2025. **Accepted:** January 14, 2026

regulatory programs across species (Magnusdottir et al. 2013; Soto and Ross 2021). Female germ cells initiate meiosis, whereas male germ cells enter the quiescent phase of the mitotic cell cycle in gonads between 70 and 80 d of fetal age (Wrobel and Suss 1998). The foundational SSC pool is established from the prespermatogonia population in calf testes during the prepubertal period (Curtis and Amann 1981). Preleptotene spermatocytes transit into the meiotic phase, and after several rounds of differentiation, full spermatogenesis is completed in the pubertal testis (Curtis and Amann 1981; Moura et al. 2011). The duration of spermatogenesis in cattle is approximately 61 d, which is comparable to the 64 d in humans (Heller and Clermont 1963). For example, although mouse SSC-enriched cultures have been successfully established (Kubota et al. 2004; Takashima and Shinohara 2018), similar conditions fail to support long-term maintenance of primary spermatogonia in bulls (Fujihara et al. 2011; Oatley et al. 2016), indicating that undefined factors function to sustain SSC activities. A comprehensive and unbiased survey of undifferentiated spermatogonial subtypes and their interactions with niches is required to fully understand the signaling pathways that direct SSC fate decisions. Furthermore, dissecting the molecular identities of spermatogenic cells and crucial factors in determining the efficiency of spermatogenesis will facilitate the discovery of new male fertility markers and the development of genetic approaches for improving bull fertility; these approaches have been largely ignored but are inevitable for increasing the efficiency of cattle production (Butler et al. 2020).

Another unexplained mystery in spermatogenesis is hybrid male sterility between closely related species in the Bovinae subfamily. These species diverged millions of years ago and contained the same number of chromosomes ($2n = 60$); however, hybridization of different members of this subfamily often causes male-specific fertility problems in F1 offspring, indicating incomplete reproductive isolation (Pathak and Kieffer 1979; MacEachern et al. 2009; Niayale et al. 2021). A well-documented case of hybrid male sterility occurs between taurine cattle and yak, two species belonging to the same genus that diverged approximately 4.9 million years ago (Qiu et al. 2012). The female offspring of male taurine cattle and female yak or the reciprocal cross are fertile, but the hybrid males are sterile because of spermatogenic arrest in the meiotic phase (Niayale et al. 2021). The Dobzhansky–Muller model proposes that hybrid incompatibilities result from interactions between at least two genes that have independently diverged in hybridizing species (Orr 1996). Meiosis defects are the major causes of hybrid sterility, and genes that play crucial roles in this process often contribute to speciation (Brideau et al. 2006; Mihola et al. 2009; Powers et al. 2020). PR domain containing 9 (Prdm9) is the first and only hybrid sterility gene identified in vertebrates (Mihola et al. 2009). Prdm9 is a meiosis-specific histone methyltransferase that directs DSB formation and recombination in mice (Baudat et al. 2010). Intersubspecific polymorphisms of the *Prdm9* allele cause errors in synapsis, recombination and sex body formation in hybrid house mice (Mihola et al. 2009; Bhattacharyya et al. 2013). Spermatogenic problems in cattle-yak appear to differ from those in hybrids of house mouse subspecies because, in addition to meiotic arrest, undifferentiated spermatogonial fate decisions are impaired (Shah et al. 2018; Zhang et al. 2024). Furthermore, similar to the phenotype observed in interspecies cat hybrids (Davis et al. 2015), spermatogenesis gradually resolves in the offspring of female cattle-yak through successive backcrossing (Tumennasan et al. 1997). Despite these findings, the identities of the spermatogenic subtypes, intricate gene expression patterns and candidate genes driving the hybrid sterility of cattle-yak remain unclear.

In this study, we performed scRNA-seq and meiotic analyses to examine the transcriptional signatures of germ cells and somatic cell subtypes of testes in cattle and yak. We then analyzed testicular cell types of cattle-yak and first-generation backcrossed offspring (BC1) and constructed developmental trajectories to determine the fate transitions of spermatogenic cells. Integrated analyses revealed a list of genes that likely play important roles in regulating spermatogonial development, meiosis, and spermiogenesis in yak and cattle. These findings indicate that defects in prophase of meiosis I are the primary cause of spermatogenic failure in cattle-yak hybrids, and that these impairments are likely due to dysregulation of subsets of genes in spermatocytes. Taken together, these results help elucidate the molecular regulatory network involved in cattle spermatogenesis and provide valuable information for exploring the mechanisms of speciation and reproductive isolation in large animals.

Results

scRNA-seq scanning of testicular cells reveals transcriptomic features of spermatogenesis in cattle and yak

First, we generated single-cell transcriptome data from freshly isolated cells from the testicular tissues of adult cattle and yak. Histological examination confirmed the occurrence of complete spermatogenesis in these samples (Figure S1a). A total of 23,522 cattle cells and 17,130 yak cells passed quality control (Table S1), and the cells of each species were visualized using uniform manifold approximation and projection (UMAP) (Figure S1b and c). Somatic cells and germ cells are recognized using specific markers, specifically *INHA* for Sertoli cells (Sohni et al. 2019), *IGF1* for Leydig cells (Shami et al. 2020), *CDH5* for endothelial cells (Stévant et al. 2018), *C1QA* for macrophages (Guo et al. 2018), *CD52* for innate lymph cells (Shami et al. 2020), *RGS5* for smooth muscle cells (Guo et al. 2021), and *DDX4* for spermatogenic cells (Sohni et al. 2019) (Fig. 1a and b; Figure S1d to f). The peritubular myoid cells were grouped together with the Leydig cells (Fig. 1a and b). An average of 1,295 genes were detected in somatic cells, and 3,929 genes were detected in the germ cells of cattle and yak (Table S1). Next, we extracted germ cells to characterize the transcriptomic patterns and subtypes of spermatogenic cells in cattle and yak. Previously known marker genes in humans (Guo et al. 2018), cynomolgus macaques (Lau et al. 2020), and mice (Green et al. 2018; Wang et al. 2019) were used to identify various spermatogenic cells (*EGR4*, *ASB9*, *MKI67*, *CRABP1*, and *DMRT1* for spermatogonia; *REC8*, *RAD51AP2*, *SYCP2*, *FKBP6*, *HORMAD2*, *AURKA*, and *CCNA1* for spermatocytes; *SPACA3*, *ACRV1*, and *LYZL1* for round spermatids; and *HEMGN, PRM2*,

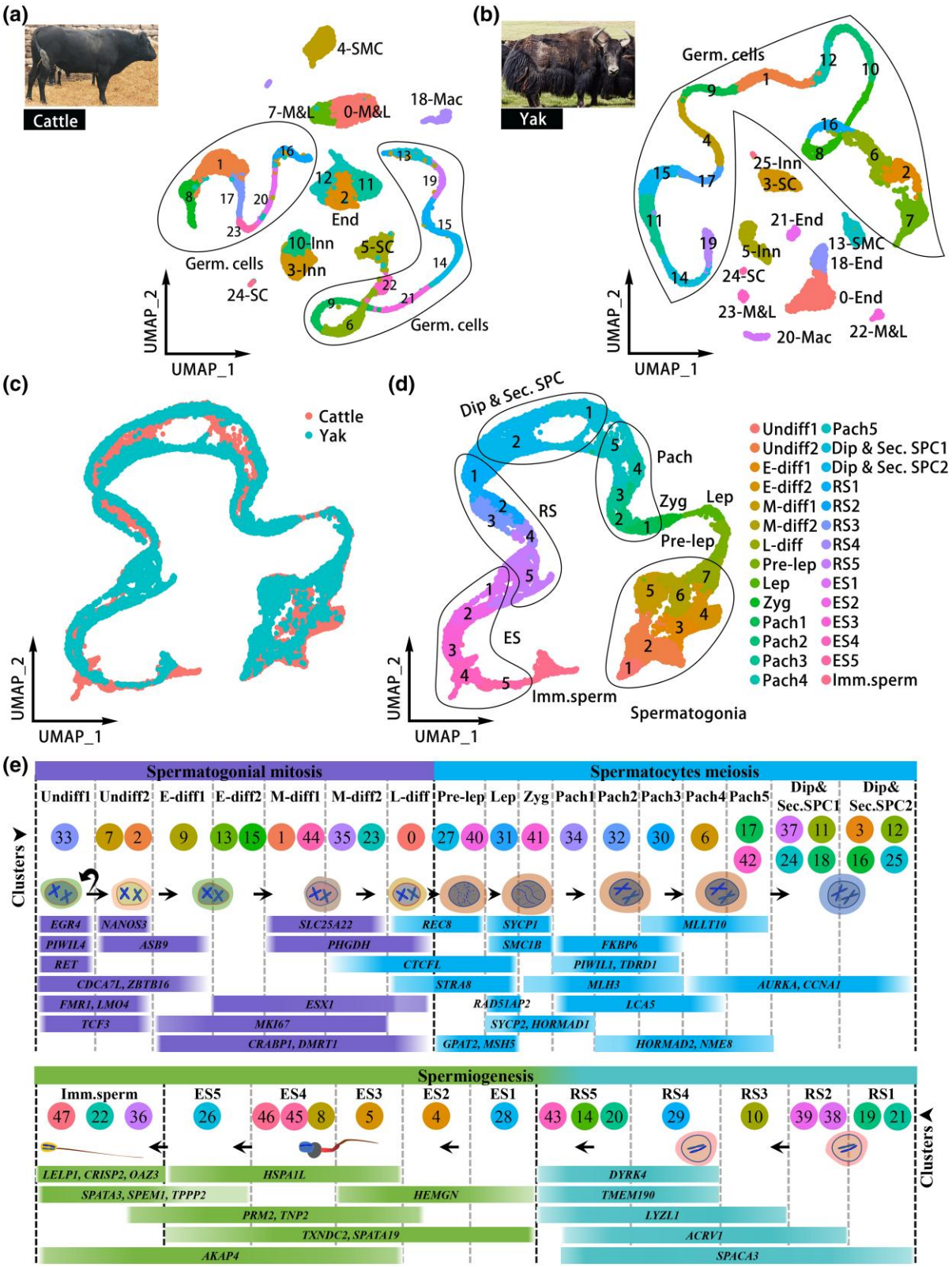

Figure 1 Overview of different cell types from scRNA-Seq analyses of yak and cattle testes. (a and b) Visualization of major cell types from the testes of yak (a) and cattle (b) in UMAP. End: Endothelial cells, Inn: Innate lymph, Mac: Macrophages, M&L: Myoid cells and Leydig cells, SC: Sertoli cells, SMC: smooth muscle cells. (c) The resulting integrated UMAP plots show the germ cell distributions from yak and cattle testes. (d) Visualization of germ cell subtypes from the testes of yak and cattle in UMAP. Prelep: preleptotene spermatocytes, Lep: leptotene, Zyg: zygotene, Pach: pachytene, Dip & Sec. SPC: diplotene spermatocytes and secondary spermatocytes, RS: round spermatids, ES: elongating spermatids, Imm.sperm: immature sperm. (e) The expression of noted marker genes in integrated germ cell clusters for cattle and yak allowed the identification of cell types. The cluster numbers were matched with the cattle and yak integrated UMAP in Figure S2a.

SPATA19, *SPATA3*, *TPPP2*, and *LELP1* for elongating spermatids) (Table S1). These analyses led to the identification of 48 clusters of spermatogenic cells in yak and cattle (Figure S2a). Pearson correlation coefficient analysis revealed that the corresponding germ cell subtypes of cattle and yak were strongly correlated (Figure S2b). Among these, two types of undifferentiated spermatogonia (Undiff1&2) and 5 types of differentiating spermatogonia in the early, middle, and late stages (E-diff1&2, M-diff1&2, and L-diff, respectively) were classified in the spermatogonial compartment. Similarly, ten distinct types of spermatocytes (preleptotene, leptotene, zygotene, pachytene 1–5, diplotene and secondary spermatocytes 1 & 2) and 11 types of spermatids (round spermatids 1–5, elongating spermatids 1–5, and immature sperm) were identified (Fig. 1d and e).

Developmental trajectory and conserved regulators of spermatogenesis in cattle and yak

We then analyzed the expression of genes enriched in each cluster (Figure S2a) and found that each spermatogenic cell subtype exhibited its own transcriptomic signature, indicating successful recognition of germ cell subpopulations (Figure S2c). Pseudotime trajectory analysis revealed that all the germ cell subtypes were arranged along the axis and generated two major branching points with five distinct states (Fig. 2a to d). A comparison of the distribution of each type of germ cell subtype over the entire differentiation trajectory revealed that the spermatogonial subtype (E-diff1) and spermatocytes at meiosis I spanned branch point 2, whereas the secondary SPC subtypes and round spermatids (RS1-5) spanned branch point 1 (Fig. 2b). We next screened gene expression dynamics over a pseudotime trajectory, and five distinct patterns were detected (Fig. 2e). The genes in group 1 were involved mainly in cellular metabolic and molecular biosynthetic processes; the genes in group 2 were associated with spermatogonial differentiation and spermatocyte meiosis, including the cell cycle process, mitotic cell cycle, cell cycle phase transition, chromosome segregation, DNA repair, double-strand break repair, and meiotic nuclear division; the genes in group 3 were involved in male gamete generation and spermatid differentiation; the genes in group 4 were related to cilium organization and movement and flagellum-dependent cell motility; and the genes in group 5 were associated with protein modification and catabolic processes (Fig. 2e). Notably, the expression of representative genes for each germ cell subtype (*ASB9*, *UCHL1*, *RADX*, and *MKI67* for spermatogonia; *REC8*, *STRA8*, *HSPA2*, and *AURKA* for spermatocytes; and *FAM170B* and *SPEM2* for spermatids) differed along the pseudotime axis (Fig. 2f). Most genes were coexpressed in various germ cell subtypes of yak and cattle (Figure S3a; Table S1), and only a small number of genes were expressed in only cattle or yak or exhibited high expression levels in yak but very low expression levels in cattle (or vice versa) (Figure S3b; Table S1). These data demonstrated that the cellular composition and developmental trajectory of cattle and yak cells were similar, indicating that the core transcriptional programs regulating the fate transition of spermatogenic cells are mostly conserved.

scRNA-Seq analyses of testicular cells from cattle-yak and backcrossed offspring

After revealing the molecular patterns of spermatogenesis in cattle and yak, we next examined the spermatogenic subtypes in cattle-yak and its first-generation backcrossed offspring (BC1). Spermatogenic defects observed in cattle-yak is gradually recovered in backcrossed progenies (Niayale et al. 2021). In this study, cattle-yak was produced by crossing female yak with a male cattle, and all BC1 animals were generated by breeding female cattle-yak with another male cattle (Fig. 3a). The testis weight significantly decreased in the cattle-yak males but increased in the BC1 males (88.22 ± 10.76 g vs 69.70 ± 14.40 g, $P < 0.001$) (Fig. 3b; Table S2). We conducted hematoxylin and eosin (H&E) staining on cross-sections of testicular tissues from cattle-yak and BC1, which were used for scRNA-seq, and found that the seminiferous tubules of cattle-yak contained only spermatogonia and a small number of spermatocytes; notably, those of BC1 contained spermatogonia, spermatocytes, and a limited number of spermatids (Fig. 3c, Figure S4a and b). The number of spermatogonia and spermatocytes significantly increased, and the number of apoptotic cells decreased in the BC1 compared with those in the cattle-yak ($P < 0.01$) (Figure S5a to d). Evaluation of the acrosome status by peanut agglutinin (PNA) staining revealed the presence of round spermatids and elongating spermatids in all BC1 testes examined; however, the density and developmental stage of the spermatids in each sample were slightly different (Figure S4c). Because spermatids were not present in the epididymis of BC1, we isolated elongating spermatids from the seminiferous tubules of these BC1 males and performed PNA staining. The results revealed that almost all the heads of elongating spermatids presented varying degrees of morphological abnormalities (Fig. 3d). We then tested the function of spermatids in BC1 testes and found that these spermatids were able to fertilize yak oocytes after microinsemination by intracytoplasmic injection, and approximately 20% of fertilized embryos developed into blastocysts (Fig. 3e to f), indicating that despite being low in number, the spermatids in the BC1 testis were functional after microinsemination.

We then generated single-cell transcriptomic profiles of testicular cells from cattle-yak and BC1 males to examine the cellular and molecular changes associated with the partial recovery of spermatogenesis. Because cattle-yak and their backcrossed offspring are hybrid animals that carry their parental genetic information, therefore we used yak and taurine cattle reference genomes to identify transcripts in the testicular cells of these two animals. There were six types of somatic cells in the cattle–yak, and a dramatic difference in the cellular composition of the germ cells was detected (Fig. 3g; Figure S5e to g). Only 11 clusters of various subpopulations of spermatogenic cells were identified (Figure S5g). An average of 5,024 genes were expressed in these germ cells (Table S2). The cells in clusters 0, 1, 2, 3, 5, 6, 8, and 10 were spermatogonia, and those in clusters 4, 7, and 9 were spermatocytes; the expression of spermatid-enriched genes, such as *SPACA3*, *ACRV1*, and *LYZL1*, did not appear in any of the clusters (Figure S5h). The genes identified from the spermatogonia were enriched in the mitotic cell cycle, cell cycle phase transition, and cellular metabolic/biosynthetic processes. The genes in the spermatocytes were related to the meiotic cell

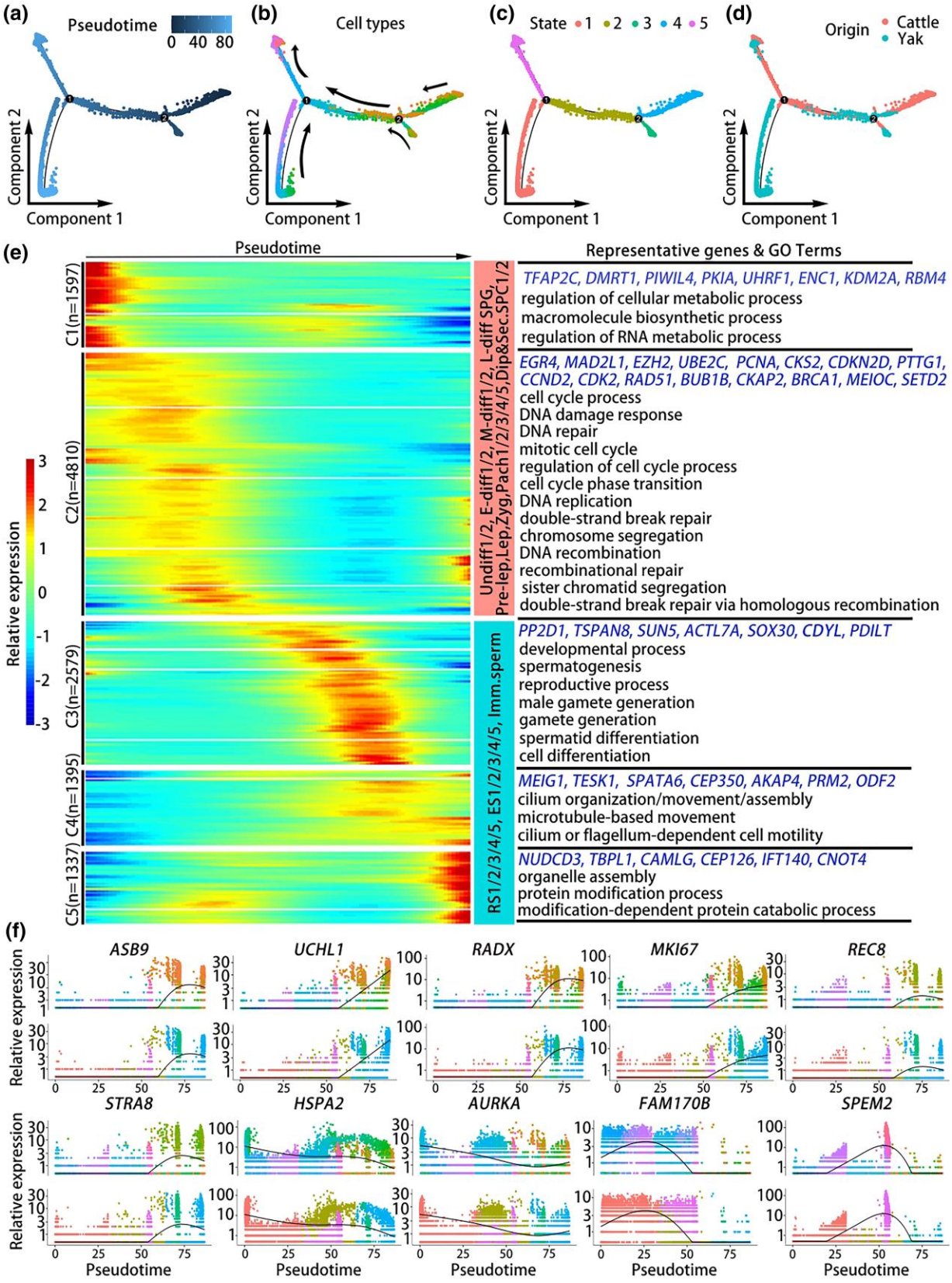

Figure 2 Gene expression dynamics and developmental trajectory of spermatogenic cells in cattle and yak. (a to c) Germ cell trajectories with cells ordered by pseudotime (a), cell type (b) and state (c) by unbiased dynamic lineage analysis. The colors of the cell types in (b) were matched to those in Fig. 1d. (d) The resulting integrated UMAP plots show the germ cell distributions from cattle and yak testes. (e) Heatmap of DEGs from different germ cell subtypes following the trajectory branch. (f) Expression patterns of key markers over pseudotime among various germ cell subtypes. The cells are colored and ordered according to the pseudotime cell type (top) and state (bottom) colors in the corresponding UMAP plots.

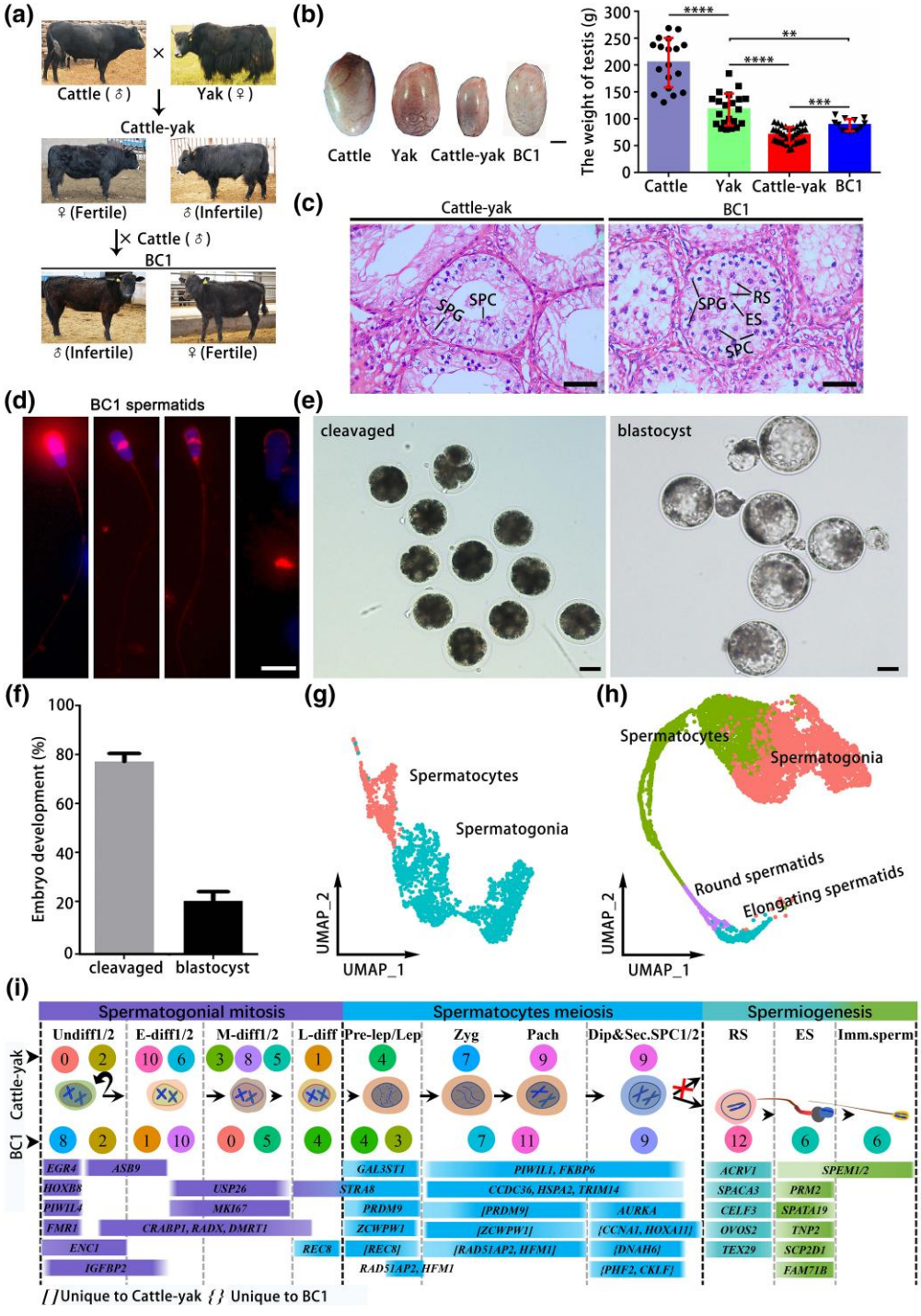

Figure 3 Analysis of the single-cell atlas of testicular tissue in the progenies of distant hybrids of cattle and yak. (a) Model diagram of cattle-yak and backcrossed progeny (BC1) produced by crossbreeding between cattle and yak. Cattle-yak were produced by breeding female yak with a male cattle. All the BC1 s were progeny of three female cattle-yak and a male cattle. (b) Testis weights of yak, cattle, cattle-yak and BC1, scale bar = 2 cm; pairwise comparisons of testicular weights of cattle ($n = 18$), yak ($n = 21$), cattle-yak ($n = 37$) and BC1 ($n = 13$) using ANOVA; (c) Hematoxylin and eosin (H&E) staining of testicular tissue from cattle-yak (left) and BC1 (right), which were from one of the samples used for scRNA-Seq (related to BC1 in Figure S4), scale bar = 50 μm. Spermatogenic cells beyond the spermatocyte stage were not visible in the cattle-yak, and spermatogenic defects were alleviated in the BC1 compared with those in the cattle-yak. SPG: spermatogonia, SPC: spermatocytes, RS: round spermatids. ES: elongating spermatids. (d) Penaut agglutinin (PNA) staining of spermatids isolated from BC1 testes. PNA staining was used to assess the acrosomal status. Scale bar =10 μm. (e and f) The function of BC1 spermatids was evaluated via intracytoplasmic injection into yak oocytes, and a proportion of cleavage-stage embryos and blastocyst-stage embryos were obtained; scale bar = 50 μm. (g–i) scRNA-Seq data were analyzed with the taurine cattle genome. (g) UMAP annotation for the germ cells of cattle-yak resulting in only two subtypes, spermatogonia and spermatocytes. (h) UMAP annotation for the germ cells of BC1 resulting in the identification of spermatogonia, spermatocytes, round spermatids, and elongating spermatids. (i) The expression of noted marker genes in clusters of cattle-yak and BC1 germ cells allowed the identification of cell types.

cycle, nuclear division, and chromosome segregation (Table S2). In total, the preliminary screening of 1,640 spermatogonia (78.58%) and 447 (21.42%) spermatocytes revealed the entire germ cell subpopulation of the cattle-yak (Table S2). These data illustrated that the fate decisions of undifferentiated spermatogonia and meiosis progression are two major defects that result in spermatogenic arrest in cattle-yak.

We then performed a similar analysis on BC1 testes. To capture all types of spermatids in the BC1 animals, three BC1 samples were pooled for scRNA-seq analysis. The transcriptome landscape of testicular cells from these animals was different from that of cattle-yak cells (Fig. 3h and i; Figure S6, Table S3). As expected, transcripts specific for elongating spermatids were present in the testes of these animals (Figure S6d and e; Table S3). Despite significant reductions in number, the seminiferous tubules of BC1 animals contained spermatogonia, spermatocytes, round spermatids and elongating spermatids, indicating partial recovery of spermatogenesis (Fig. 4). These cells were distributed into 13 clusters (Figure S6c), and an average of 4,666 genes were expressed in each cell (Table S3). Like those in adult yak and cattle, the germ cells in 12-month-old BC1 exhibited a continuous developmental trajectory (Figures S6e, and 7a and b). The testes of 12-month-old BC1 contained 21.40% and 8.38% spermatocytes and spermatids, whereas those of aged matched cattle-yak contained only 0.87% spermatocytes (Figure S7c and d). These data indicated that the subtypes of germ cells in 12-month-old BC1 cattle were similar to those in adult yak and cattle. Furthermore, we analyzed the data for cattle-yak and BC1 using the yak genome, and the same cell composition and detectable expressed genes were used for subsequent multicombination comparisons (Figure S8; Tables S2 and S3). Overall, we found that spermatogenic defects were partially recovered in BC1 animals, although the overall number of each spermatogenic cell type was lower than that in cattle or yak, and the gene expression of each spermatogenic cell was slightly different.

Detection of synapsis, DSB repair and reductional meiosis recombination frequencies in cattle-yak

Although it has been known for years that spermatocyte development is blocked in cattle-yak, a detailed analysis of meiosis progression is still lacking. Next, we used spermatocyte nuclear spreading across meiotic stages, which involves chromosome synapsis and homologous recombination, to detect the process of meiosis prophase I (Zickler and Kleckner 2015). The staining of SYCP1 and SYCP3, components of the synaptonemal complex, revealed the presence of spermatocytes at the leptotene, zygotene, pachytene, diplotene, and diakinesis stages in the cattle-yak (Fig. 5a, Figure S9a). The ratio of pachytene spermatocytes decreased, but leptotene, diplotene, and diakineses spermatocytes increased in cattle-yak, indicating abnormal meiosis progression (Fig. 5b). Furthermore, quantitative analysis revealed that the percentage of pachytene spermatocytes with normal synapsis was lower in cattle-yak than that in yak ($86.17 \pm 2.27\%$ vs $97.07 \pm 0.82\%$, $P < 0001$), indicating defects in pachytene synapsis (Figure S9b). We then investigated whether the formation or repair of DSBs was impaired in the cattle-yak. During meiosis in male mammals with normal fertility, homologous recombination is initiated from DSBs, which start from the leptotene of meiosis I, go through zygotene and pachytene spermatocytes and are marked by γH2AX (Zakharyevich et al. 2010). DSBs are catalyzed by the endonuclease SPO11, and their subsequent repair involves the recombinases RAD51 and DMC1 (Cloud et al. 2012). All DSBs in autosomal chromosomes are completely repaired in pachytene spermatocytes; however, the DSBs in sex chromosomes are not repaired until the diplotene phase of spermatocytes (Cloud et al. 2012). γH2AX staining was first detected in leptotene spermatocytes of both yak and cattle-yak (Fig. 5d), and a comparable number of RAD51 foci were present in leptotene spermatocytes of yak and cattle-yak (Fig. 5c), indicating that DSBs were properly initiated and then started the repair process at this stage. In yak, only a few RAD51 foci were present on the XY chromosomes of pachytene spermatocytes; however, RAD51 foci remained on autosomes in cattle-yak at this stage (Fig. 5c and f). We found that the X and Y chromosomes remained unsynapsed in yak and cattle-yak spermatocytes (Fig. 5a). As expected, γH2AX staining was detected only in sex bodies in pachytene spermatocytes of yak; however, it was detected on autosomes, and the staining signal on sex bodies was absent in pachytene spermatocytes of cattle-yak, indicating that the formation of sex bodies was impaired (Fig. 5d and g). These data suggest that the DSB repair process and the formation of sex bodies in pachytene spermatocytes are abnormal and incomplete in cattle-yak (Fig. 5e and h).

The presence of spermatids in the BC1 testis indicated that the DSBs were at least partially repaired. Indeed, the number of RAD51 foci in BC1 was significantly lower than that in cattle-yak ($P < 0.05$) (Fig. 5i and j). Moreover, a certain number of DSBs remained in autosomes for some pachytene spermatocytes in BC1, but the sex bodies formed in ∼50% of the pachytene spermatocytes, and a much lower number of γH2AX foci were present in BC1 spermatocytes than in spermatocytes of cattle-yak (Fig. 5k and l), indicating that a greater proportion of spermatocytes formed the sex body. Interestingly, the formation of spermatids was independent of the PRDM9 genotype in BC1 (Figure S9b). *Prdm9* is responsible for hybrid male sterility because of its asymmetric binding in mice (Davies et al. 2016), and its zinc finger (ZF) array directs the positions of DSBs by binding to specific DNA sequences (Grey et al. 2011). The *PRDM9* gene is also polymorphic in cattle (Ahlawat et al. 2017), and we detected six and five ZF domains of *PRDM9* in cattle and yak, respectively. The composition of germ cells in the seminiferous tubules of BC1 was not related to the number of ZF arrays (Figure S9c to d). Together, these data suggested that the meiosis error was partially recovered in BC1 animals and that a novel genetic mechanism likely causes hybrid sterility in cattle-yak.

Differential gene expression in undifferentiated spermatogonia associated with hybrid sterility

Considering that the proliferation and differentiation of undifferentiated spermatogonia is the first and most fundamental event in spermatogenesis, we extracted spermatogonia and conducted integrated analyses to elucidate differences in the fate

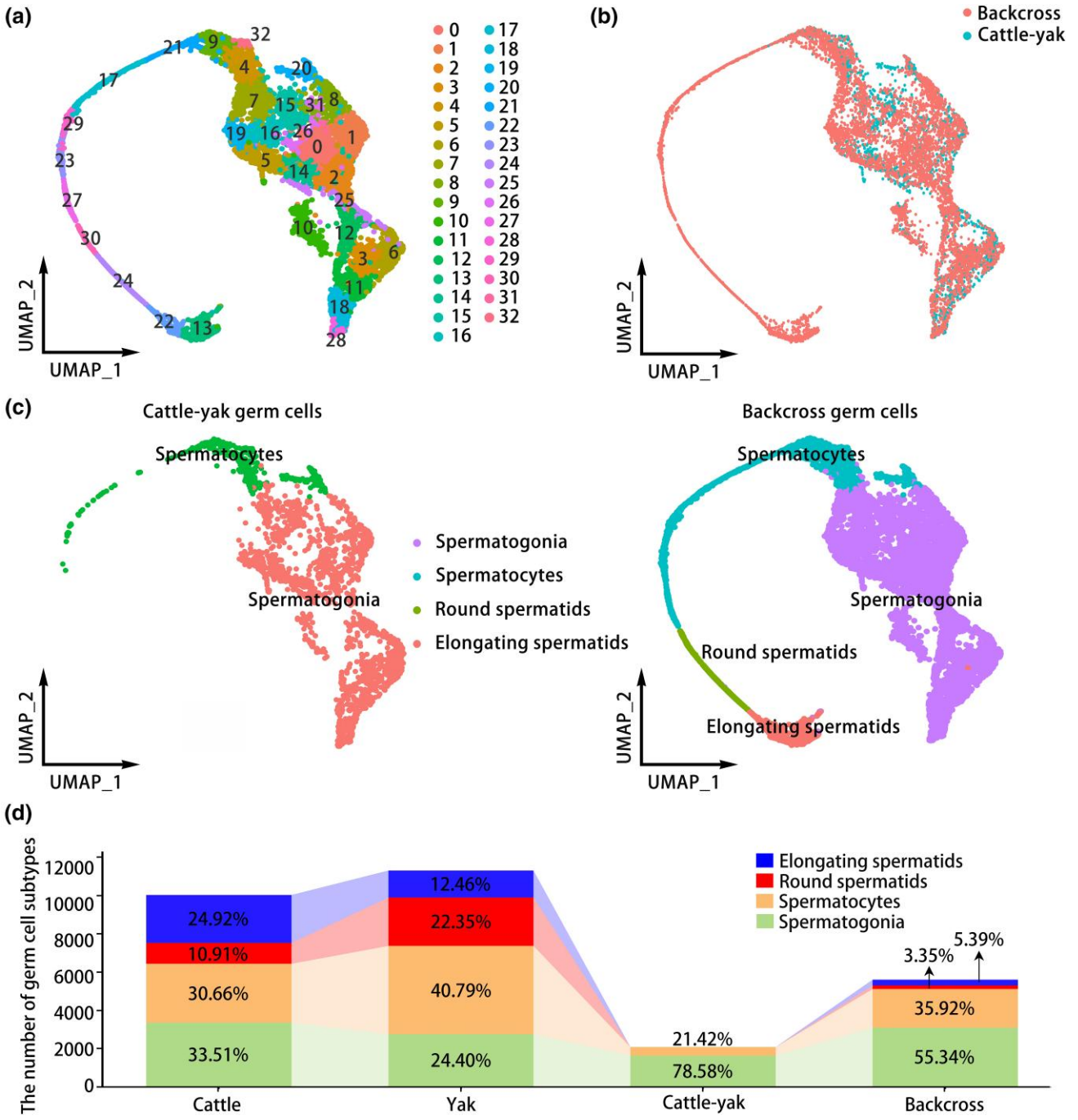

Figure 4 Integrated analysis of germ cells extracted from cattle-yak and BC1 testes based on the taurine cattle genome. (a) Unbiased clustering of germ cells extracted from cattle-yak and BC1 in UMAP plots. (b) The resulting integrated UMAP plots show the germ cell distributions of the cattle-yak and BC1 testes. (c) UMAP annotation for the germ cells of cattle-yak resulting in only two subtypes, spermatogonia and spermatocytes, while the germ cells of BC1 were spermatogonia, spermatocytes, round spermatids, and elongating spermatids. (d) Quantification of the number of germ cells in cattle, yak, cattle-yak, and BC1.

transitions and developmental trajectories of yak, cattle, and their hybrid progenies. The yak genome and taurine cattle genome were used for both cattle-yak and BC1. In total, spermatogonia (yak: 2,446; cattle: 3,061; cattle-yak: 1,452 (taurine cattle genome), 1,591 (yak genome); BC1: 2,498 (taurine cattle genome), 2,780 (yak genome)) were extracted, and these cells were divided into 25 clusters (Figure S10a; Table S4) and further assigned to two different undifferentiated spermatogonial clusters (Undiff1 and Undiff2), two early differentiated spermatogonial clusters (E-diff1 and E-diff2), two middle differentiated spermatogonial clusters (M-diff1 and M-diff2), and one late differentiated spermatogonial cluster (L-diff1) (Fig. 6a; Figure S10b–d). The numbers of all seven subtypes of spermatogonia in cattle and yak were equivalent. In contrast, virtually every spermatogonial subtype in cattle-yak was less abundant, and the reduction in cell number was alleviated in BC1 (Fig. 6b; Table S4).

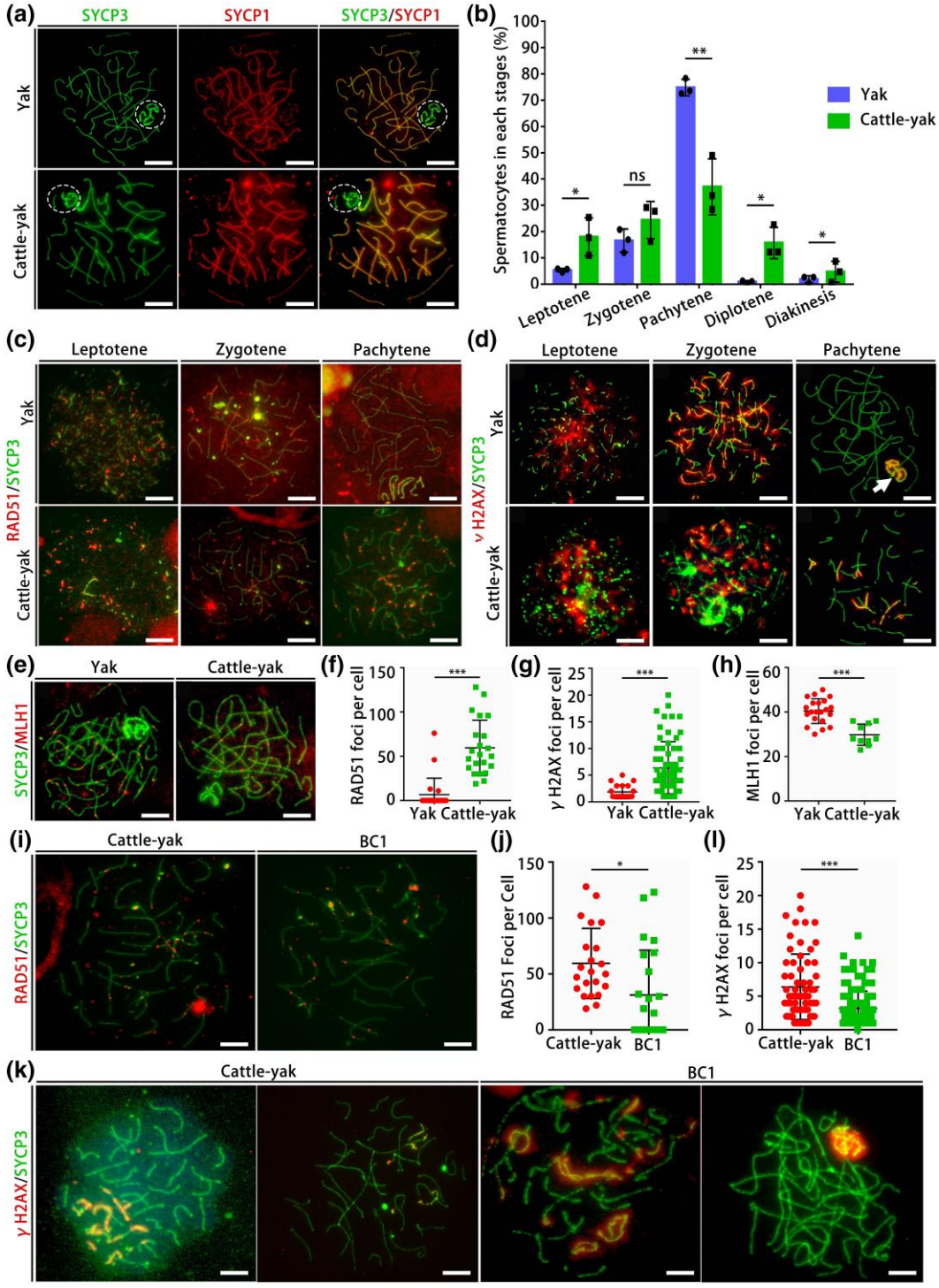

Figure 5 Analysis of meiotic events in spermatocytes from yak, cattle-yak, and BC1 testes. (a) Spermatocytes of yak and cattle-yak were labeled with SYCP1 (red) and SYCP3 (green). Dashed circles, unsynapsed X and Y chromosomes. Scale bar = 5 μm. (b) The proportions of yak and cattle-yak spermatocytes at each stage of meiosis I prophase, including leptotene, zygotene, pachytene, diplotene, and diakinesis, $P < 0.05$*; $P < 0.01$**, ns, not significant. $n = 100$. (c) Detection of the DSB repair process conducted by RAD51 localization analysis in spermatocytes of yak and cattle-yak. Scale bar = 5 μm. (d) γH2AX location analysis in leptotene, zygotene, and pachytene spermatocytes of yak and cattle-yak. The white arrow indicates sex body marked by γH2AX. Scale bar = 5 μm. (e) Detection of recombination frequency by MLH1 location analysis in pachytene spermatocytes from yak and cattle-yak. Scale bar = 5 μm. (f) Statistics of RAD51 foci in pachytene spermatocytes from yak and cattle-yak. $P < 0.001$***. $n > 20$. (g) The number of γH2AX foci in pachytene spermatocytes from yak and cattle-yak. $P < 0.001$***. $n > 50$. (h) The number of MLH1 foci in pachytene spermatocytes in yak and cattle-yak. $P < 0.001$***. $n > 10$. (i) Detection of RAD51 localization in pachytene spermatocytes from BC1. Scale bar = 5 μm. (j) Number of RAD51 foci in pachytene spermatocytes from BC1 and cattle-yak. $P < 0.05$*. $n > 20$. (k) γH2AX location analysis in pachytene spermatocytes from BC1. Scale bar = 5 μm. (l) Statistics of γH2AX foci in pachytene spermatocytes from BC1 and cattle-yak. $P < 0.001$***. $n > 50$. Each experiment was conducted on samples from at least three animals ($n = 3$).

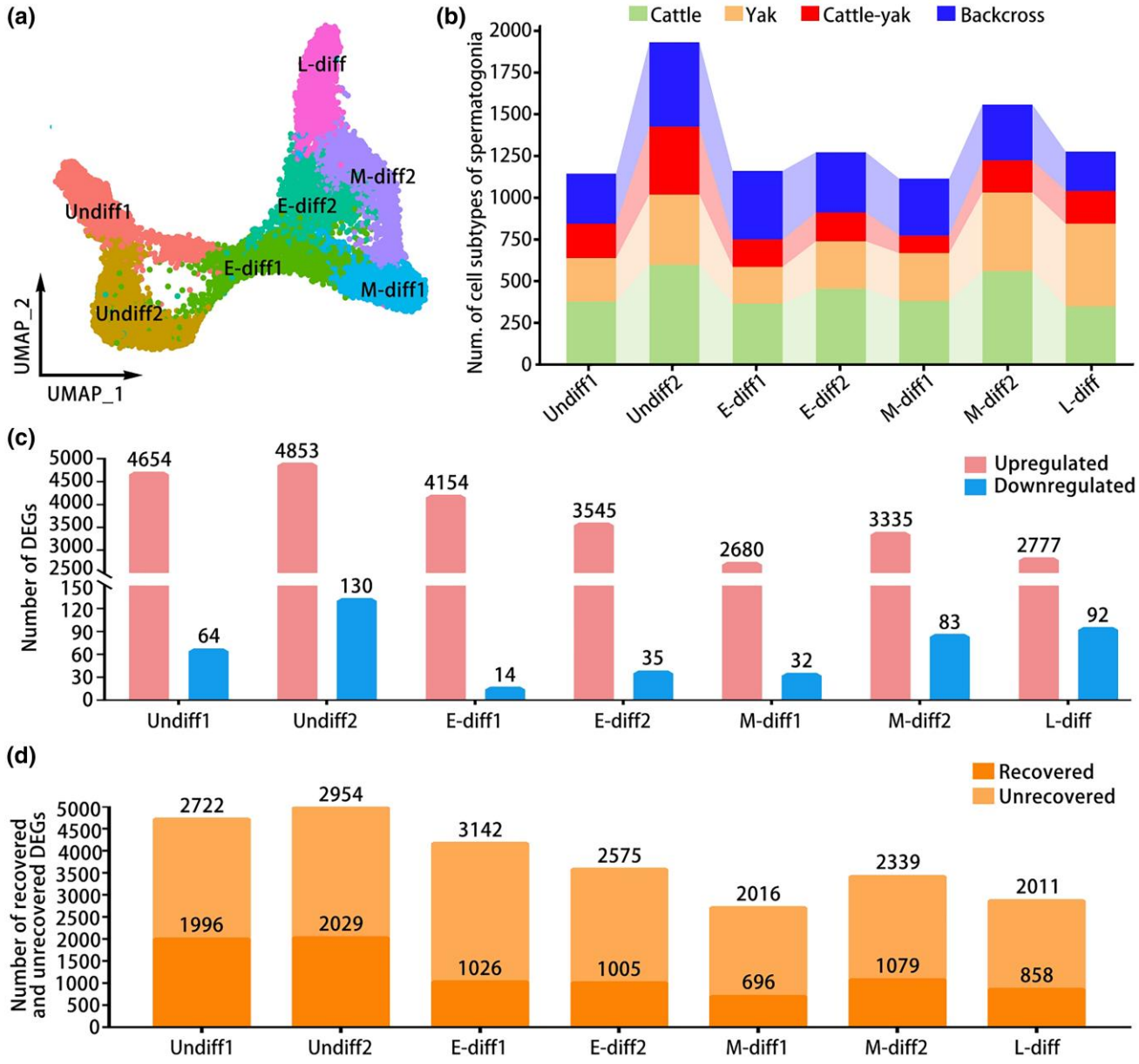

Figure 6 Screening of differentially expressed genes in various subtypes of spermatogonia from cattle, yak, cattle-yak, and BC1. (a) Integrated annotation of spermatogonia from cattle, yak, cattle-yak, and BC1. Cattle-yak and BC1 data were analyzed using both the yak genome and the taurine cattle genome. (b) Histogram showing the number of spermatogonial subtypes isolated from cattle, yak, cattle-yak, and BC1. (c) Histogram showing the DEGs in various spermatogonial subtypes in cattle-yak compared with yak and cattle. (d) Number of genes recovered or not recovered in spermatogonial subtypes of the BC1 testes.

Next, based on the taurine cattle genome for cattle-yak and BC1, we integrated the cells of cattle, yak, and cattle-yak to detect genes that potentially determine the imbalanced fate determination within subtypes of undifferentiated spermatogonia. Compared with those in cattle and yak, a total of 5,843 differentially expressed genes (DEGs) were identified in the spermatogonia of cattle-yak (Table S4). We hypothesized that genes whose expression was up- or downregulated in the cattle-yak but returned to normal levels in BC1 likely had a role in hybrid sterility. Therefore, we screened genes that exhibited such expression patterns, and as a result, 3,917 DEGs were identified by comparing the cells from cattle-yak and BC1 (Fig. 6c; Table S4). Two undifferentiated spermatogonial subtypes (Undiff1 and Undiff2) from cattle-yak exhibited dramatic differences in gene expression (Fig. 6c; Table S4). In total, 2,930 genes (36.51%) in cattle-yak returned to normal levels in BC1, whereas 5,095 genes (63.49%) in cattle-yak still presented abnormal expression levels in BC1 (Fig. 6c and d; Table S4). Both types of undifferentiated spermatogonia had greater recovery ratios (42.41% and 40.72%) than the average recovery ratio of all the spermatogonia subtypes (32.42%). Notably, we identified a list of genes whose expression was downregulated in cattle-yak but upregulated in BC1 in various spermatogonial subtypes (Figure S10d; Table S4).

Similarly, we also used the same methods and screened the dysregulated genes in cattle-yak but recovered in BC1 based

using the yak genome as a reference. A total of 5,310 DEGs were identified in the spermatogonia of cattle-yak, and 3,960 DEGs were identified by comparing the cells from cattle-yak and BC1. In total, 3,259 genes (42.23%) in cattle-yak returned to normal levels in BC1, whereas 4,458 genes (57.77%) in cattle-yak still presented abnormal expression levels in BC1. Both types of undifferentiated spermatogonia had greater recovery ratios (56.49% and 52.72%) than the average recovery ratio of all the spermatogonia subtypes (38.62%). Furthermore, when the two genomes of cattle-yak and BC1 were compared, 1,986 genes were recovered from BC1 (Table S4). These analyses revealed candidate genes that have potential roles in determining spermatogonial fate decisions in hybrid animals.

Screening genes associated with meiotic defects in spermatocytes of cattle-yak

Although differences in cell fate decisions and gene expression exist in spermatogonia, these cells can still enter meiosis and form primary spermatocytes. We conducted an integrated analysis of spermatocytes and screened genes that may be involved in meiotic arrest. To this end, we extracted spermatocytes and further characterized the molecular identities of the spermatocytes, among which cattle-yak and BC1 were analyzed using taurine cattle and yak genomes (Fig. 7a; Figure S11a and b; Table S5). The overall number of spermatocytes (255 in taurine cattle and 257 in yak genomes) in cattle-yak was less than 10% of that in cattle (2,998) and yak (4,566). The spermatocyte population increased more than 4-fold (1,360 in the taurine cattle genome and 1,297 in the yak genome) in BC1, indicating partial recovery of spermatocyte arrest (Fig. 7b; Table S5). In the cattle-yak, the ratios of preleptotene spermatocytes to diplotene spermatocytes and secondary spermatocytes 1 decreased, but advanced secondary spermatocytes were missing (Fig. 7b). These cells reappeared in the BC1 testes; the number of all spermatocyte subtypes was greater than that in the cattle-yak testes (Fig. 7b).

A DEG analysis of cattle-yak and BC1 revealed 1,947 upregulated and 1,430 downregulated genes in the spermatocytes of cattle-yak compared with those of yak and cattle (Table S5). We subsequently examined DEGs in spermatocytes of BC1 and identified 1,632 upregulated and 1,025 downregulated genes compared with those in cattle-yak (Table S5). These genes were assigned to different spermatocyte subtypes: 121 genes in preleptotene, 13 in leptotene, 26 in zygotene, 307 in pachytene spermatocytes, and 1,300 in Dip & Sec. SPC1 spermatocytes returned to normal in BC1 (40.14%), whereas 2,310 genes (59.86%) did not change (Fig. 7c and d; Table S5). Compared with the average recovery ratio of all spermatocyte subtypes (25.11%), the recovery ratio of SPC1 spermatocytes was greater (28.14% and 61.21%).

We also employed the same methods and screened the genes dysregulated in cattle-yak but recovered in BC1 based on the yak genome. A total of 3,284 DEGs were identified in the spermatocytes of cattle-yak, and 2,338 DEGs were identified by comparing the cells from cattle-yak and BC1. In total, 1,830 genes (45.55%) in cattle-yak returned to normal levels in BC1; interestingly, 2,188 genes (54.45%) in cattle-yak still displayed abnormal expression levels in BC1. Compared with the average recovery ratio of all the spermatocyte subtypes (34.11%), the recovery ratio of SPC1 spermatocytes was greater (44.32% and 68.72%). Furthermore, in the comparison of the two genomes for cattle-yak and BC1, 1,233 genes were recovered in BC1 (Table S5). Notably, the transcripts recovered in preleptotene of BC1 were involved in metabolism and biosynthetic processes, and those recovered in pachytene and diplotene spermatocytes were closely related to the meiosis process, including nuclear division, metaphase/anaphase transition of the cell cycle, chromosome segregation, double-strand break repair, and the DNA damage response and repair (Figure S11c and d; Table S5).

Collectively, these data revealed that the dysregulation of genes in spermatocytes of cattle-yak became more severe with the progression of meiosis, whereas genes with greater recovery ratios were present in BC1. Together with the results of the nuclear spreading analysis, we concluded that meiosis is blocked at the late diplotene and secondary spermatocyte stages in F1 cattle-yak hybrids and that a small proportion of spermatocytes develop throughout the whole process to form spermatids in BC1.

Disrupted meiotic sex chromosome inactivation and gene upregulation on sex chromosomes are important factors in hybrid male sterility in mice, and the underlying mechanism is also present in cats (Davis et al. 2015). Therefore, we tested 1,056 X-linked genes and found that 41 genes were highly expressed in pachytene, diplotene and secondary spermatocytes in cattle-yak. X-linked genes constituted 13.36% of all upregulated genes in the pachytene spermatocytes of cattle-yak. Interestingly, the expression levels of these genes in these BC1 cells were comparable to those in yak and cattle, indicating that they returned to normal levels (Fig. 8; Table S6). These data further illustrated the existence of abnormal meiotic sex chromosome inactivation in spermatocytes of cattle-yak.

Screening genes associated with impaired spermiogenesis in BC1 males

Next, we examined the molecular identities of spermatids in BC1 testes and attempted to screen candidate genes causing defective spermiogenesis in these animals. Spermatids were extracted from yak (3,933 cells), cattle (3,576 cells) and BC1 (435 from the taurine cattle genome and 442 from the yak genome), and different subtypes of spermatids, including five round spermatid subtypes (RS1, RS2, RS3, RS4, and RS5), five elongating spermatid subtypes (ES1, ES2, ES3, ES4, and ES5), and immature sperm, were identified (Fig. 7e). The proportions of various subtypes of spermatids decreased in BC1 (Fig. 7f, Figure S12a and b; Table S7). A total of 1,544 DEGs were detected in BC1 compared with yak and cattle, and 1,213 DEGs were screened in BC1 using the yak reference genome (Table S7). The majority of DEGs existed in round and early elongating spermatids, with a significant proportion of upregulated genes in BC1, indicating the impairment of transcription programs that negatively control gene expression (Fig. 7g; Table S7). In total, in the comparison of the two genomes for BC1, 106 upregulated and 189 downregulated genes in round spermatids and 90 upregulated and 127 downregulated genes in elongating spermatids were screened in BC1 (Table S7).

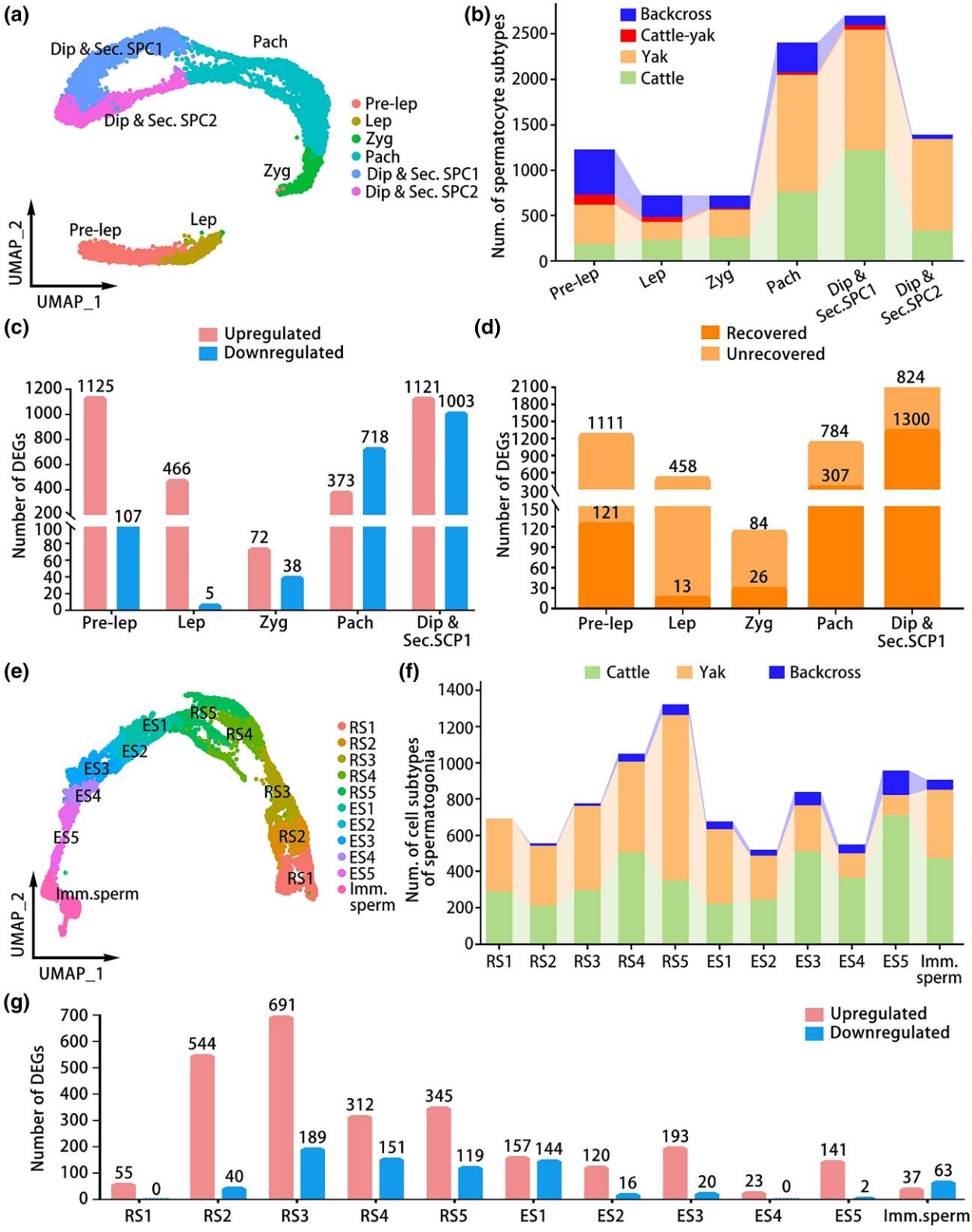

Figure 7 Integrated annotation and screening of differentially expressed genes in spermatocytes and spermatids. (a) Integrated annotation of spermatocytes obtained from cattle, yak, cattle-yak, and BC1, among which cattle-yak and BC1 were analyzed using both the yak genome and the taurine cattle genome. (b) Histogram showing the number of spermatocyte subtypes isolated from cattle, yak, cattle-yak, and BC1. (c) Histogram showing the DEGs in various spermatocyte subtypes in cattle-yak compared with yak and cattle. (d) The number of genes recovered or not recovered in the spermatocyte subtypes of BC1. (e) Integrated annotation of spermatids obtained from yak, cattle, and BC1. (f) Histogram showing the number of spermatid subtypes isolated from yak, cattle, and BC1. (g) Histogram showing the DEGs in various spermatid subtypes in BC1 compared with those in yak and cattle.

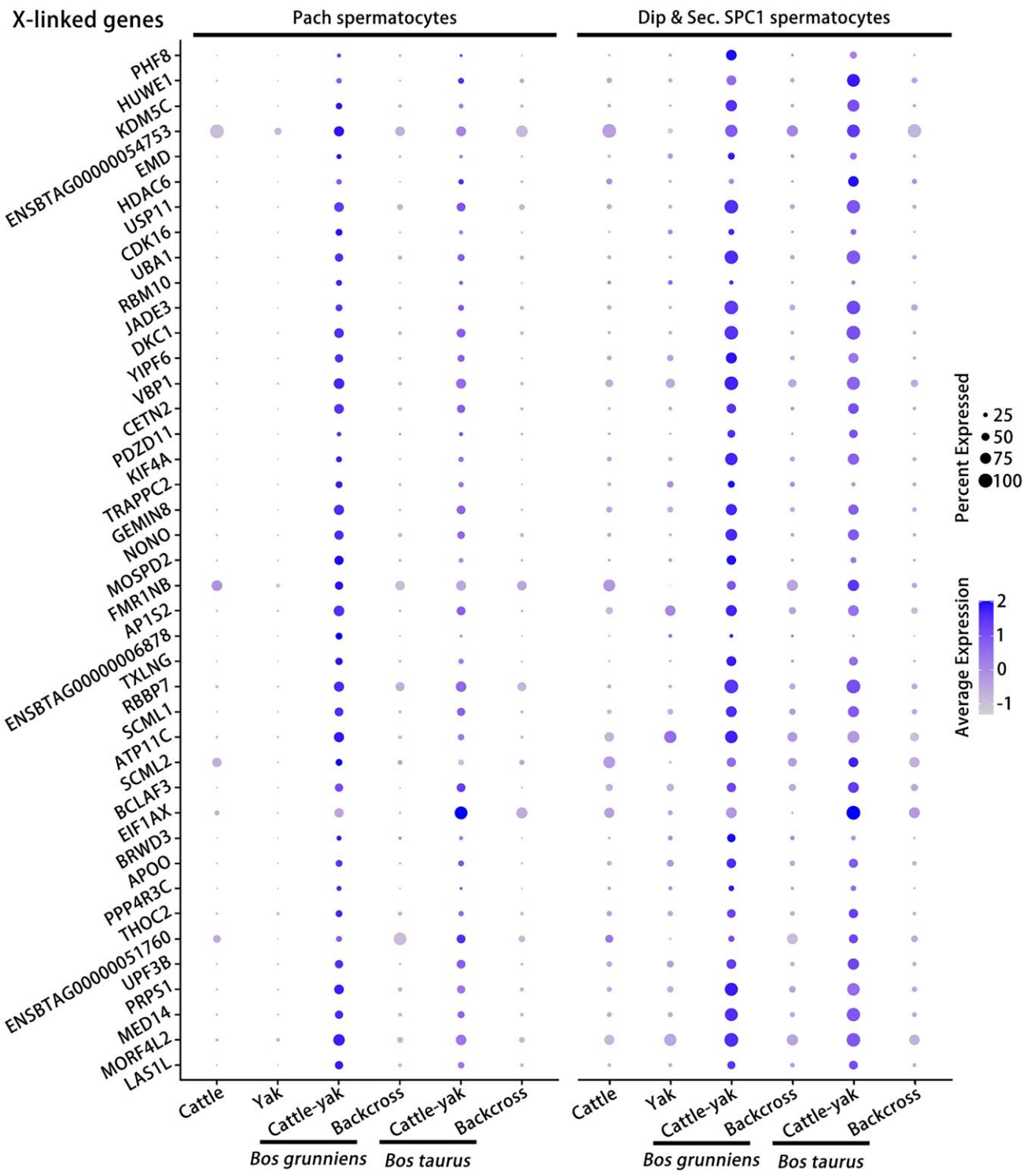

Figure 8 Dot plot of screened X-linked genes with relatively high expression levels in spermatocytes of cattle-yak. The results were analyzed by using both the yak and the taurine cattle genomes from cattle-yak and BC1. Pach: pachytene spermatocytes, Dip: diplotene spermatocytes, Sec. SPC: secondary spermatocytes.

The DEGs implicated in this critical process for spermiogenesis were further elucidated through ClueGO analysis (Figure S12c). This analysis revealed that genes involved in the regulation of spermatid development and differentiation, sperm capacitation, dephosphorylation and flagellated sperm motility were all altered in the spermatids of BC1 animals. Together, these data revealed that although spermatids were formed in early backcrossed males, these spermatids were rare in number, and the gene expression signatures did not fully recover to normal levels in cattle or yak. These data provide important information for further examination of candidate genes involved in hybrid sterility related to spermatid differentiation.

Genomic structural variations in differentially expressed genes related to hybrid sterility in cattle and yak

Hybrid incompatibilities caused by genomic structural variations (SVs) have been identified in other species (Shen et al. 2017). SVs widely exist in the yak and taurine cattle genomes, and these genetic differences play important roles in yak adaptation and domestication (Zhang et al. 2021; Gao et al. 2022). During meiosis, abnormal DNA replication, break repair, and recombination are associated with SV formation (Segura-Wang et al. 2017). We focused on the genes whose expression was up- or downregulated in the spermatogenic cells of the cattle-yak testes but returned to normal levels in BC1 compared with that in the cattle and yak (using the yak genome or the taurine cattle genome). We next asked whether these restored DEGs carried SVs and identified a total of 661 genes that contained SVs in intronic, exonic, or untranslated regions or 150 bp upstream and downstream flanking regions (Table S8). Among these genes, 528 were upregulated and only 133 were downregulated in the spermatogenic cells of the cattle-yak. These SVs were associated with upregulated genes in spermatogonia and downregulated genes in spermatocytes. Specifically, 85.42% (451) of the upregulated genes were assigned to spermatogonia, whereas 100% (133) of the downregulated genes were detected in spermatocytes (Fig. 9a; Table S8). More than 95% of these SVs were in intronic or upstream and downstream regions of genes, including *TDRD1* in spermatocytes (Fig. 9b; Table S8). Similarly, we found that 48 and 36 SV harboring DEGs in round and elongating spermatids were associated with SVs (Fig. 9c). These data suggest that SVs in the regulatory regions of genes are likely associated with gene expression in the spermatogenic cells of cattle-yak and BC1.

To reveal the associations between SVs and differential gene expression in the cattle-yak testis, we first analyzed bulk RNA-seq data from testicular tissues of cattle, yak, and cattle-yak ($n = 6$ for each) and identified a total of 7,741 DEGs from 23,158 genes in the testes of cattle-yak. Previously, we reported that 6,733 highly selected SV genes are signatures associated with the genetic divergence of cattle and yak (Gao et al. 2022) and further extracted 6,204 genes that were expressed in the testes of taurine cattle, yak, cattle-yak, and BC1. Using these two datasets, we conducted Fisher's exact test (two-sided), and the results revealed that a greater proportion of genes containing SVs were differentially expressed, indicating a significant association between SVs and gene expression ($P = 1.75 \times 10^{-5}$, Table S9). Next, we tested whether DEGs were associated with SVs in different spermatogenic subpopulations using scRNA-seq data from cattle, yak, cattle-yak and BC1 testes (Table S10). SVs and DEGs were associated in undifferentiated spermatogonia, differentiated spermatogonia, and secondary spermatocytes ($P < 0.05$, Table S10). Finally, because meiosis arrest is the main cause of spermatogenic failure in cattle-yak and this impairment was partially reversed in BC1, we analyzed the restored DEGs that exhibited differential protein abundance and further asked whether these DEGs contained SVs. To this end, we compared the lists of restored DEGs with differentially abundant proteins (DAPs) in spermatocytes of cattle-yak (Wu et al. 2023) and identified 115 genes whose mRNA and protein expression were dysregulated in spermatocytes of cattle-yak. These genes are involved in piRNA processing, transposable element silencing, chromosome condensation, P granule organization, and other biological processes related to spermiogenesis (Fig. 9d; Table S11). Among these DEGs, 24 genes carrying SVs have known functions in meiosis and spermiogenesis, including *TDRD1*, *CLGN*, CDK5RAP2, and *DDHD1* (Fig. 9e; Table S11). These results revealed a potential relationship between SV and disordered DNA repair and reduced recombination frequency during meiosis in cattle-yak.

Discussion

Hybrid sterility between closely related species serves as a reproductive barrier and often occurs in male offspring of mammals following Haldane's rule (Coyne 1985). Hybrid male sterility (HMS) is predominantly caused by meiotic arrest, which results in the death of spermatocytes and sterility. Defects in meiosis have been documented in hybrids between blue foxes and silver foxes (Gustavsson et al. 1988), horses and donkeys (Yang et al. 2004), and interspecies felines (Davis et al. 2015). Genetically, hybrid incompatibility is caused by genomic variations arising from evolution between parental species, including differences in the sequence, structure, and location of protein-coding genes or noncoding DNA (Lemmon and Kirkpatrick 2006). Although dissecting differential gene expression in the testes of hybrids and their parents is insufficient to reveal the genetic mechanisms of HMS because most DEGs are a consequence of spermatogenic arrest, not the direct cause of sterility, understanding the cell composition, developmental trajectory of spermatogenic cells, and cell-specific changes in gene expression in hybrids is crucial (Runemark et al. 2025). Therefore, exploring the transcriptomic landscape of various spermatogenic cells is a prerequisite for identifying hybrid male sterility factors (Civetta 2016). In this study, we investigated the cellular composition and gene expression dynamics of spermatogenesis in cattle, yak, cattle-yak, and BC1. We explored intricate gene expression patterns related to hybrid sterility in cattle-yak. Because the X chromosome is proposed to play a major role in the reproductive isolation of closely related mammalian species (Larson et al. 2016), we examined the expression of X chromosome-linked genes in the spermatocytes of cattle-yak hybrids. Combined with histological and meiotic progression analysis, we showed that meiosis failure is the primary cause of hybrid sterility in cattle-yak and that this change is partially reversed in backcrossed progeny. We also identified 24 genes carrying genomic SVs that potentially play a role in the hybrid incompatibility between cattle and yak and likely other species from the Bovinae subfamily.

Identifying testicular cell subtypes and corresponding gene expression signatures is indispensable for obtaining a comprehensive understanding of spermatogenesis. Using data from mice (Green et al. 2018), humans (Sohni et al. 2019; Guo et al. 2021), and nonhuman primates (Lau et al. 2020) as references, we accurately recognized all testicular somatic lineages and spermatogenic cells. We assigned Sertoli cells, macrophages, endothelial cells, peritubular myoid cells, Leydig cells, innate lymph cells, smooth muscle cells, and various subtypes of germ cells. In total, 11 subtypes of spermatogonia, ten subtypes of spermatocytes,

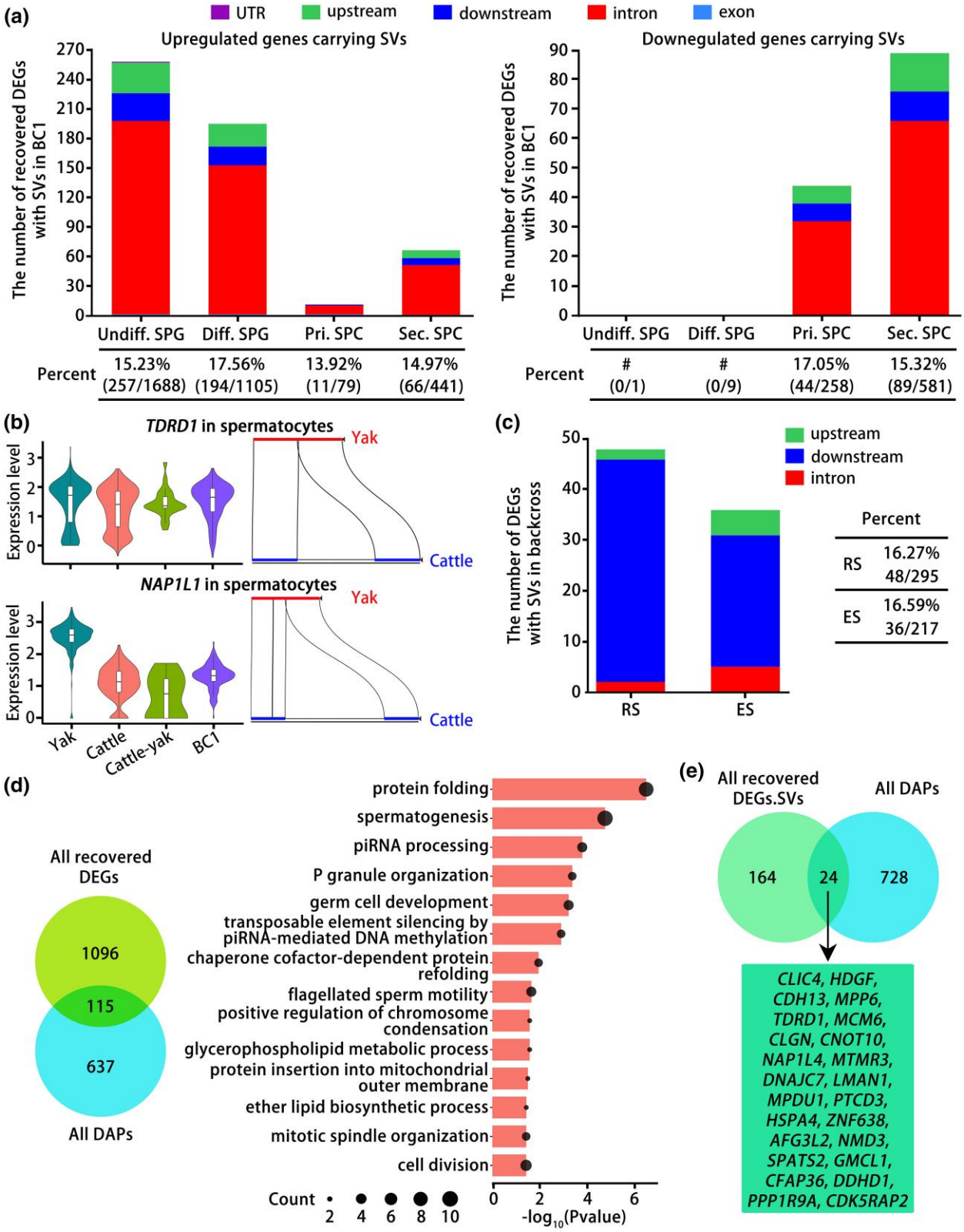

Figure 9 Screening of DEGs with structural variants in various germ subtypes of cattle, yak, cattle-yak, and BC1. The results were obtained by using the yak and taurine cattle reference genomes for cattle–yak and BC1. (a) Histogram showing the number of upregulated and downregulated genes in cattle-yak but recovered in BC1, with structural variants in various subpopulations of spermatogonia and spermatocytes. (b) Schematic diagram of the violin plots and structural variants of representative genes. (c) Histogram showing the number of DEGs with structural variants in various spermatid subtypes of BC1 compared with yak and cattle. (d) Venn diagrams and GO enrichment analysis of all DEGs with restored expression levels in spermatocytes of BC1 and DAPs in spermatocytes of cattle-yak. (e) Venn diagrams of all recovered DEGs. SVs in spermatocytes of BC1 and all DAPs in spermatocytes of cattle-yak. DAPS: differentially abundant proteins.

five subtypes of round spermatids, and five subtypes of elongating spermatids and immature sperm were identified for both yak and cattle. Although testicular single-cell transcriptomics of cattle-yak has been reported recently (Mipam et al. 2023; Wang et al. 2023), these studies did not examine germ cells from taurine cattle and lacked detailed analyses to reveal the molecular signatures of different spermatogenic cells. We identified a list of genes that were exclusively expressed in only one or several germ cell subtypes. These genes were expressed in various germ cell subtypes ranging from undifferentiated spermatogonia to immature sperm. Most genes were shared by cattle and yak, but a minority of genes were expressed only in cattle or yak. These results indicate that conservation and divergence coexist in the germ cells of cattle and yak. Species-specific genes may have important functions in the genomic incompatibility that causes hybrid sterility.

The hybrid sterility of cattle-yak and backcrossed progeny is a complex phenotype that involves defects in all three phases of spermatogenesis. Spermatocytes are the most advanced germ cells in the seminiferous tubules of cattle (Niayale et al. 2021). Spermatids are present in the testes of backcrossed animals, but a detailed analysis is lacking (Zhao et al. 2023). In this study, we examined the molecular identities of spermatogenic cells in BC1 testes, which partially recover spermatogenesis. We used testicular samples from 12-month-old BC1, similar to those from the testes of adult yak and cattle, and spermatogenic cells from BC1 at this developmental stage included spermatogonia, spermatocytes, and spermatids. Compared with yak and cattle, the number of each subtype of spermatocytes was extremely low, and meiosis was completely arrested at the diplotene and diakinesis stages. The number of pachytene spermatocytes was also low and abnormal synapsis was more frequent in cattle-yak, but we speculated that a subset of pachytene spermatocytes passed this check point before being completely arrested at the final stage of meiosis I. Interestingly, the numbers and subtypes of meiotic cells increased, and a proportion of spermatocytes successfully underwent meiosis and entered the early stage of spermiogenesis process in BC1 testes. We detected the presence of elongating spermatids in the testes of BC1, which can serve as a resource to screen candidate genes that guide meiosis progression and direct the early stage of spermatid development. Notably, although the three BC1 s used in scRNA-seq are derived from the same male parent, substantial differences in genetic composition still exist among them. Therefore, although spermatocytes in each BC1 sample successfully completed the meiosis process and further developed into spermatids, the density, and developmental stage of the spermatids in each BC1 sample slightly differed.

The genes that were dysregulated in the cattle-yak but recovered in BC1 were related to all aspects of meiosis, including chromatin condensation, chromosome synapsis, initiation and repair of DSBs, homologous recombination, crossovers formation, metaphase chromosome segregation, and spindle stability. During mammalian meiosis, the prophase of meiosis I is a remarkable event, as the key genetic material exchange process–chromosome synapsis and homologous recombination occur in a successive order during preleptotene, leptotene, zygotene, pachytene, diplotene, and diakinesis (Lie et al. 2009). There is a strong relationship between the collapse of spermatogenesis and the disruption of synapsis and abnormal homologous recombination (Mu et al. 2014; Zickler and Kleckner 2015; Floriot et al. 2021). Interestingly, unlike what was observed in a hybrid house mouse line (Davies et al. 2016), only approximately 13% of pachytene spermatocytes in cattle-yak exhibited abnormal synapsis. Meiotic arrest coincided with DSB retention in pachytene spermatocytes, a lower homologous recombination ratio, and the absence of γH2AX-marked sex bodies in cattle-yak hybrids. Consistent with these findings, the expression of X-linked genes was upregulated in spermatocytes of these F1 hybrids. Because the disruption of meiotic sex chromosome inactivation and the overexpression of X-linked genes are strongly associated and constitute the key molecular mechanism that causes hybrid sterility (Turner 2015), we propose that genes that are dysregulated but recovered in BC1, especially those located on the X chromosome, are strong candidates for further investigation of the underlying mechanisms of meiotic failure and spermatid development.

Screening genetic variations between cattle and yak is a crucial step in identifying genes dictating reproductive isolation in the Bovinae subfamily. The Dobzhansky–Muller model proposes that hybrid sterility between closely related species is caused by genetic incompatibilities of divergent alleles (Orr 1996). Previous works have established crucial roles of single-nucleotide polymorphisms and cis-regulatory elements in regulating hybrid viability or sterility (Mack and Nachman 2017; You et al. 2023). Recent work has shown that SVs have profound effects on various biological processes (Alkan et al. 2011) and that these types of large-scale mutations can also impact speciation (Stuart et al. 2023). The deleterious effects of SVs on the expression of speciation genes have been verified in hybrid male sterility in plants (Shen et al. 2017). These SVs can change gene expression and cause defects in meiotic recombination, therefore resulting in failure of proper gametogenesis. These findings suggest that SVs between yak and cattle may cause deleterious consequences in the spermatogenesis of hybrid offspring, especially in the meiotic phase. In this study, we showed that SVs and differential gene expression were significantly associated. A number of SV-harboring genes were dysregulated in the spermatogonia and spermatocytes of cattle-yak but were restored to normal expression levels in the cells of BC1 males, indicating their possible involvement in spermatogenic recovery. The expression patterns of these genes were completely opposite in spermatogonia and spermatocytes. Specifically, most of these SV-carrying genes were upregulated in the spermatogonia but downregulated in the spermatocytes of cattle- yak. Integrated analysis with comparative proteomic data resulted in the identification of 24 candidate genes causing spermatogenic arrest, especially meiotic problems in cattle and yak. Dissecting the functional role of sequential variations in meiotic genes is essential for understanding the genetic basis of meiotic recombination (Johnston 2024). Therefore, all these DEGs carrying SVs constitute a dataset for subsequent screening for key molecules involved in abnormal spermatogonial differentiation and meiotic arrest in cattle-yak and aberrant spermiogenesis in BC1 males. Because the expression of the majority of the DEGs identified in this study is likely caused by the incompatibility of causative genes in hybrid animals, the mechanisms underlying the expression and functions of these SV genes in spermatogenic cells await further investigation.

Materials and methods

Animals

All animal experiments were performed in accordance with the Guide for the Care and Use of Laboratory Animals and were approved by the Animal Ethics and Welfare Committee at the Northwest Institute of Plateau Biology, Chinese Academy of Sciences (approval code: hwipb012). Testicular tissues were obtained under local anesthesia according to a previously described approach (Li et al. 2020). Testis weights were calculated from data from 18 cattle, 21 yak, 37 cattle-yak, and 13 BC1 individuals (Table S2). Freshly collected testicular samples from cattle ($n = 3$, 3 years old), yak ($n = 4$, 3 years old), cattle-yak ($n = 5$, 3 years old), and backcrossed offspring (BC1, $n = 3$, 1 year old) were used for scRNA-seq analysis. The cattle-yak were produced by crossing five different female yak with a male cattle, and the BC1 animals were generated by breeding three female cattle-yak with another cattle bull. Briefly, the caudal epididymis and tunica vaginalis were trimmed and washed with PBS three times. Fresh testicular samples were cut into small pieces and then treated as follows: one part was used to prepare a single-cell suspension for scRNA-seq; one part was placed in 4 °C storage medium (70% Dulbecco's modified Eagle's medium/F12 (DMEM/F12), 20% FBS, and 10% dimethyl sulfoxide) for 2 h, then −20 °C for 2 h, −80 °C overnight, and finally transferred to liquid nitrogen for further analysis (Zhang et al. 2013); furthermore, other slices were immobilized in 4% paraformaldehyde (PFA) for immunohistochemistry and immunofluorescence and Boun's fluid for morphological analysis. Finally, testicular tissues from three cattle, four yaks, five cattle-yaks, or three backcrossed offspring were pooled for library construction and scRNA-seq, respectively. Notably, each sample used for scRNA-seq was processed separately for histological and cytological analyses.

Sample digestion for single-cell RNA sequencing

Testicular samples from yak, cattle, cattle-yak, and BC1 males were collected to prepare single-cell suspensions by taking appropriate slices from specific species. Single-cell suspensions of testicular tissue were prepared by a two-step enzyme digestion method. Briefly, fresh testicular tissue was washed three times with Dulbecco's phosphate-buffered saline (DPBS) and then digested with 1 mg/mL collagenase type IV (Sigma–Aldrich) at 37 °C until the seminiferous tubules became loose for approximately 20 min. After natural sedimentation of the digestive mixture on ice, the supernatant was removed. The remaining tissues were washed once with DPBS and further digested by adding 0.05% trypsin–0.04% ethylene diamine tetraacetic acid (TE, Sigma–Aldrich) and 1 mg/mL hyaluronidase (Sigma–Aldrich) for 30 min. The digestion was terminated with medium (Dulbecco's modified Eagle's medium/nutrient mixture F-12 (DMEM/F12) containing 10% FBS) after most tubules were completely digested. After filtration using a 70 μm sieve and collection by centrifugation at 400 × g for 8 min, the cells were resuspended and then treated with red blood cell lysis buffer (Solarbio Life Sciences, R1010, Beijing, China) to remove red blood cells, followed by washing with DPBS two times. After that, the cells were resuspended in cold sample buffer provided by the Single Cell Capture and cDNA Synthesis Kit (Cat. No. 633731) within the BD RhapsodyTM single-cell analysis system and used for single-cell library construction and sequencing followed by counting with 27200 cells, and the percentage of viable cells was >85% for every sample.

scRNA-seq library construction and sequencing

Each sample was subjected to single-cell 3′ whole-transcriptome amplification (WTA) (Doc ID: 210966) to construct a single-cell library in a BD Rhapsody system (Becton, Dickinson and Company, New Jersey, USA). The cell suspension was loaded into a special U-shaped groove using a BD Rhapsody Cartridge Kit (Cat. No. 633733), and the cells naturally settled into the micropores. Then, excess magnetic beads with cell labels and unique molecular identifier (UMI) sequences were loaded to ensure that the vast majority of micropores fell into the magnetic beads, which captured the mRNA released after cell lysis. After the cell capture beads were retrieved and washed several times, standard sequencing libraries were constructed through reverse transcription with the help of a BD RhapsodyTM cDNA Kit (Cat. No. 633773) and cDNA amplification using a BD RhapsodyTM WTA Amplification Kit (Cat. No. 633801). Finally, the libraries were sequenced on an Illumina NovaSeq sequencer with 150 bp paired-end reads at Novogene Bioinformatics Technology Co., Ltd. (Tianjing, China).

scRNA-seq data analysis

In accordance with the BD Rhapsody pipeline, the generated raw scRNA-seq data were processed for a quality check via fastp (v0.23.0) (Chen et al. 2018), and the raw reads of Read1 and Read2, which later contained cell labels and UMI and gene information, respectively, were qualified for subsequent analysis. The reference genomes of taurine cattle (ARS-UCD1.2, GenBank: GCA_002263795.2; gene annotation file: Bos_taurus.ARS-UCD1.2.102.gff3.gz) were used for cattle, cattle-yak and BC1 (Rosen et al. 2020), whereas a newly assembled domestic yak reference genome (NWIPB_DYAK_1.0, GenBank: GCA_027580245.1) generated in our previous study was used for yak, cattle-yak, and BC1 (Gao et al. 2022). The index of the genome was established via genome files and annotation files, and the reads were mapped to the reference genome by STAR (Spliced Transcripts Alignment to a Reference) (v2.7.1a, https://github.com/alexdobin/STAR) (Dobin et al. 2013). With the default parameters (Smith et al. 2017), the reads were analyzed via UMI tools and a pipeline (https://umi-tools.readthedocs.io/en/latest/index.html) to obtain cell count matrices after cell identification, barcode extraction, read filtration and assignability to genes, and molecular counting. We removed the potential doublets of each sample using scDblFinder package (v1.12.0) and then performed subsequent analysis as described previously (Germain et al. 2020).

Seurat analysis

The generated matrices were processed by using the Seurat package (v3.2.0, https://github.com/satijalab/seurat) (Butler

et al. 2018; Stuart et al. 2019), and the lower-quality cells with the characteristics of genes expressed in fewer than three single cells, no more than 200 genes expressed in a cell, and cells with more than 5% mitochondrial sequence were discarded. The generated data were normalized through "LogNormalize," "FindVariableFeatures" was used to calculate highly variable genes, and the threshold was set at 2000. For the subsequent principal component analysis and cluster analysis, 25 principal components (PCs) were chosen based on elbow plots and stable variance, and check_duplicates were set to FALSE for nonlinear dimension reduction (uniform manifold approximation and projection for dimension reduction, UMAP) analysis (Becht et al. 2019), which reduces the cell distribution to two-dimensional space for downstream analysis. The clusters were identified using "FindClusters" with the parameter "resolution = 0.6" for single sample analysis based on clustree package (v0.5.0), and the data structures were visualized by UMAP. The genes within clusters were identified through "FindAllMarkers followed by marker gene screening with the parameters "min.pct = 0.25", "logfc.-threshold = 0.25," and "p_val_adj < 0.05."

Cell type identification and functional enrichment analysis

The relationship between each cluster and cell type was established according to the cluster-enriched genes and known marker genes for testicular germ cells and somatic cells in the published literature. The DEGs within the same cell types between cattle, yak, cattle-yak, and BC1 were identified using the "FindMarker" function in the Seurat toolkit (v3.2.0, https://github.com/satijalab/seurat) with the following parameters: "min.pct = 0.25," "logfc.threshold = 0.25," and "p_val_adj <0.05." For the Seurat toolkit with the Wilcoxon test and Bonferroni correction, the screening parameters were "min.pct = 0.25," "logfc.threshold = 0.25," and "padj_val < 0.05." The g:Profiler website (https://biit.cs.ut.ee/gprofiler/gost) was used for GO enrichment analysis, and the built-in Benjamini–Hochberg method was used for correction. $P_{adj} < 0.05$ was used as the criterion for significance screening.

Integration analysis of diverse cell subtypes

For multisample integrative analysis, germ cells were extracted from cattle and yak and then integrated for exploration of normal germ cell development with complete spermatogenesis. The Pearson correlation coefficient between the yak and taurine cattle germ cell subtypes was analyzed via the corr.test function in the psych package (v2.4.6). The spermatogonia were also extracted (cattle, yak, cattle-yak, and BC1) and integrated to further classify the spermatogonial subtypes and DEGs in hybrid offspring of yak. All spermatocytes were extracted and integrated for identification of the subtypes and DEGs in cattle-yak and BC1. Similarly, the spermatids extracted from cattle, yak and BC1 were then grouped, and the DEGs in BC1 were screened. Each germ cell subtype for both cattle-yak and BC1 was extracted based on the results of the reference genomes of *Bos taurus* and *Bos grunniens*. Data integration for each germ cell subtype from two or more samples was performed via reciprocal principal component analysis (Song et al. 2023). The input data were the list of Seurat objects and then used by the "FindIntegrationAnchors" function in the Seurat toolkit (v4.2.1, https://github.com/satijalab/seurat) to identify the anchors, which were then passed to the IntegrateData function in the Seurat toolkit and returned to the Seurat object. The new assay for multiple samples thus resulted from the integration of the corresponding expression matrices (parameter "dims = 1:25"). Cell cluster analysis was performed again on the basis of the UMAP method (Becht et al. 2019), and the parameters in FindClusters were resolution = 3.2 for the integration of spermatogonia, spermatocytes and spermatids.

Cell trajectory analysis

Cell pseudotime analysis for single germ cells was used to detect the cell differentiation trajectory and the sequence of the cell development process according to different cluster, state, and pseudotime values. The Monocle 2 package (v2.10.1) (Qiu et al. 2017) was used in the analysis of the developmental trajectories of the germ cells. DEGs were identified and further subjected to pseudotime analysis with a threshold of p_val_adj <0.01. DEGs were stored in Seurat objects, which were then reduced to two dimensions via the discriminative dimensionality reduction with trees (DDRTree) method (Trapnell et al. 2014). Heatmaps for the genes used in the pseudotime analysis were generated by using the plot_pseudotime_heatmap function in the Seurat toolkit (v3.2.0, https://github.com/satijalab/seurat).

Genome-wide identification of structural variations

SV analysis was performed as described previously (Sedlazeck et al. 2018). Briefly, the genomes of a female taurine cattle (Bos_taurus.ARS-UCD1.2) and the Y chromosome of a male taurine cattle (Btau_4.6.1) were merged into the taurine genomic file. Third-generation sequencing data from 3 taurine cattle and 26 yak samples (accession number PRJNA540974) were aligned with the ARS-UCD1.2 genome via NGMLR (v0.2.7, https://github.com/philres/ngmlr), followed by sorting and indexing with SAMtools (v1.8, https://github.com/samtools/samtools) (Danecek et al. 2021). The program cuteSV (v1.0.11, https://github.com/tjiangHIT/cuteSV) was used to call SVs for each individual with parameters (–max_cluster_bias_INS 100 –diff_ratio_merging_INS 0.3 –max_cluster_bias_DEL 100 –diff_ratio_merging_DEL 0.3) for the nanopore data or parameters (–max_cluster_bias_INS 100 –diff_ratio_merging_INS 0.3 –max_cluster_bias_DEL 200 –diff_ratio_merging_DEL 0.5) for the PacBio data (Jiang et al. 2020). The VCF files were merged using the "50 1 1 –1 –1 –1 –1" parameter of SURVIVOR (v1.0.7, https://github.com/fritzsedlazeck/SURVIVOR) software (Jeffares et al. 2017). Each sample at all SV breakpoints was genotyped, and all mapping results were mapped via cuteSV software with the option "–genotype." The data were subsequently filtered with the parameters (–max-missing 0.5–mac 3–minQ 30) of vcftools (v0.1.17, https://vcftools.github.io/) to remove unreliable SVs.

Finally, the data were annotated with VEP (https://asia.ensembl.org/index.html?redirect=no). The FST (fixation index) values were calculated via vcftools (v0.1.17) (Danecek et al. 2011), and the SV with a high FST value (top 0.5%) was used as a candidate SV for downstream analysis as previously described (Gao et al. 2022). The results revealed that 6,733 genes carried SVs (Gao et al. 2022). Furthermore, we screened for genes expressed in the testes of yak and taurine cattle that carried SV, resulting in the "SV gene" list containing 6,204 genes (Table S9).

Association analysis between SVs and differential gene expression

To examine the associations between SVs and differential gene expression using bulk RNA-seq data, testicular RNA-seq data were downloaded from the National Center for Biotechnology Information ($n = 6$ yak, $n = 6$ cattle, and $n = 6$ cattle-yak) (Table S9). The data were aligned via HISAT2 software (v2.2.1) and then normalized via the featurecounts program of subread software (v2.0.6) to obtain the expression matrix. The matrix was analyzed with the R package DEseq2 (v1.34.0) for differential expression analysis. Genes with a $P_{adj} < 0.05$ and an absolute value of log2FoldChange ≥ 1 were considered DEGs. We detected 7741 DEGs in cattle-yak compared with both yak and cattle and further obtained a list of overlapping genes (DEGs.SV, 2211) compared with 6204 SV genes (Table S9). We also integrated the yak reference genome and the taurine cattle reference genome to obtain a background set (all genes, 23158), which were expressed in the testes of taurine cattle, yak, cattle-yak, and BC1 (Table S9). Association analysis between DEGs in the testicular transcriptome and the SV gene was performed via Fisher's exact test.

To examine the relationships between SVs and differential gene expression in various spermatogenic populations using scRNA-seq data, DEGs identified from various germ cell subtypes, including the ones DEGs that were dysregulated in cattle-yak but restored in undifferentiated spermatogonia; differentiated spermatogonia, primary spermatocytes, and secondary spermatocytes of BC1; and DEGs in round and elongated spermatids of BC1 compared with those in yak and cattle (Table S10). To explore the relationships between these DEGs in cattle and the 6204 SV genes, associations were calculated via Fisher's exact test (two-sided), and the *P* values were corrected for multiple testing using the test (two-sided) Benjamin–Hochberg method. To examine the overlap between DEGs and differential protein abundance, 1,096 genes that were dysregulated in spermatocytes of cattle-yak but restored in those of BC1 were screened in a previously published proteomic dataset, which included 637 differentially DAPs in spermatocytes of cattle-yak (Wu et al. 2023).

Chromosome spreading across meiotic stages

Spermatocyte nuclear spreads were carried out as previously described (Li et al. 2020). The frozen testicular tissue was thawed in a water bath at 37 °C, cut into pieces and resuspended after being washed three times with DPBS. The tissue was left standing on ice for 3 min, and 1 mL of the supernatant was used for centrifugation to collect the cell precipitates. After resuspension in 500 μL of normal saline, the cell suspension was reacted in 500 μL of hypotonic solution (30 mM Tris-Cl, 50 mM sucrose, 17 mM sodium citrate, and 5 mM EDTA in ddH_2O) at room temperature (RT) for 20 min. After centrifugation at 5,000 rpm for 4 min, the ruptured cell precipitates were resuspended in 200 μL of 100 mM sucrose solution. A 20 μL cell suspension was evenly spread on slides soaked in 1% PFA and placed in a hot and wet box. After 2 h, the lid was lifted, and the sample was slowly dried. The slides were washed twice with 0.2% PhotoFlo200 once per minute, placed in antibody dilution buffer (ADB, 1% BSA, 0.1% gelatin, and 0.5% Triton X-100) for 1 h and dried at RT. The primary antibody was diluted with ADB, dropped onto slides, sealed with rubber cement, and incubated at 37 °C for 24 h. After removing the cement, the slides were immersed in PBS for 15 min to automatically detach the cover glass. The cells were soaked in ADB for 1 h and dried, and then the secondary antibody was added and incubated at 37 °C for 2 h. After washing with PBS for 1 h, the slices were stained with H33342, mounted in 50% glycerin, and examined under a fluorescence microscope (Lecia, Germany). The information on the antibodies used in this study is listed in Table S12.

Statistics and reproducibility

Statistical analysis of testis weights for cattle, cattle-yak, yak, and BC1 was performed by two-tailed ANOVA with Tukey's multiple comparisons test; the confidence interval was 95%. All quantitative data from at least three independent biological replicates of testicular cells were counted, and two-tailed *t* tests were used to statistically determine the differences between the means. The significance levels were set to $^{*}P < 0.05$, $^{**}P < 0.01$, and $^{***}P < 0.001$. All the data are shown as the means ± SDs and were analyzed with GraphPad Prism software (v7.01).

Author contributions

Q.Y. designed and led the project. S.W., R.W., R.Y., X.G., and G.J. analyzed the single-cell RNA-sequencing data. G.W. performed chromosome spreading, PCR and sequencing of the PRDM9 zinc finger array, oocyte maturation, fertilization, and embryo culture with the assistance of R.Z.Q.Y. and Y.F. provided the samples. S.W. and X.G. constructed the BD single-cell library. S.W., G.W., R.W., and Q.Y. drafted the paper, which was edited by all the authors.

Supplementary material

Supplementary Figures, Tables, and Methods are available at *Molecular Biology and Evolution* online.

Funding

This research was funded by grants from the National Key R&D Program (2021YFD1200405), the National Natural Science

Foundation of China (31972574 & U22A20447), the Natural Science Foundation of Qinghai (2025-ZJ-970T), and High-tech Industrialization Special Fund of Jilin and CAS (2024SYHZ0003). Q.E.Y. was supported by the "Kun Talents" project from the Qinghai Provincial Government.

Conflicts of interest

The authors declare no conflicts of interest.

Data availability

We have deposited the testicular single-cell RNA-sequencing data of domestic yak, taurine cattle, cattle-yak and backcrossed hybrids in the Sequence Read Archive database of the National Center for Biotechnology Information database with accession BioProject codes: PRJNA929587.

References

Ahlawat S *et al.* Evolutionary dynamics of meiotic recombination hotspots regulator PRDM9 in bovids. *Mol Genet Genomics*. 2017:292:117–131. https://doi.org/10.1007/s00438-016-1260-6.

Alkan C, Coe BP, Eichler EE. Genome structural variation discovery and genotyping. *Nat Rev Genet*. 2011:12:363–376. https://doi.org/10.1038/nrg2958.

Baudat F *et al.* PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science*. 2010:327:836–840. https://doi.org/10.1126/science.1183439.

Becht E *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*. 2019:37:38–44. https://doi.org/10.1038/nbt.4314.

Bhattacharyya T *et al.* Mechanistic basis of infertility of mouse intersubspecific hybrids. *Proc Natl Acad Sci U S A.* 2013:110:E468–E477. https://doi.org/10.1073/pnas.1219126110.

Brideau NJ *et al.* Two dobzhansky-muller genes interact to cause hybrid lethality in *Drosophila*. *Science*. 2006:314:1292–1295. https://doi.org/10.1126/science.1133953.

Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol*. 2018:36:411–420. https://doi.org/10.1038/nbt.4096.

Butler ML, Bormann JM, Weaber RL, Grieger DM, Rolf MM. Selection for bull fertility: a review. *Transl Anim Sci*. 2020:4:423–441. https://doi.org/10.1093/tas/txz174.

Charlier C *et al.* NGS-based reverse genetic screen for common embryonic lethal mutations compromising fertility in livestock. *Genome Res.* 2016:26:1333–1341. https://doi.org/10.1101/gr.207076.116.

Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018:34:i884–i890. https://doi.org/10.1093/bioinformatics/bty560.

Civetta A. Misregulation of gene expression and sterility in interspecies hybrids: causal links and alternative hypotheses. *J Mol Evol*. 2016:82:176–182. https://doi.org/10.1007/s00239-016-9734-z.

Cloud V, Chan YL, Grubb J, Budke B, Bishop DK. Rad51 is an accessory factor for Dmc1-mediated joint molecule formation during meiosis. *Science*. 2012:337:1222–1225. https://doi.org/10.1126/science.1219379.

Coyne JA. The genetic basis of Haldane's rule. *Nature*. 1985:314:736–738. https://doi.org/10.1038/314736a0.

Curtis SK, Amann RP. Testicular development and establishment of spermatogenesis in Holstein bulls. *J Anim Sci.* 1981:53:1645–1657. https://doi.org/10.2527/jas1982.5361645x.

Danecek P *et al.* The variant call format and VCFtools. *Bioinformatics*. 2011:27:2156–2158. https://doi.org/10.1093/bioinformatics/btr330.

Danecek P *et al.* Twelve years of SAMtools and BCFtools. *Gigascience*. 2021:10:giab008. https://doi.org/10.1093/gigascience/giab008.

Davies B *et al.* Re-engineering the zinc fingers of PRDM9 reverses hybrid sterility in mice. *Nature*. 2016:530:171–176. https://doi.org/10.1038/nature16931.

Davis BW *et al.* Mechanisms underlying mammalian hybrid sterility in two feline interspecies models. *Mol Biol Evol.* 2015:32:2534–2546. https://doi.org/10.1093/molbev/msv124.

de Rooij DG. The nature and dynamics of spermatogonial stem cells. *Development*. 2017:144:3022–3030. https://doi.org/10.1242/dev.146571.

Di Persio S *et al.* Single-cell RNA-Seq unravels alterations of the human spermatogonial stem cell compartment in patients with impaired spermatogenesis. *Cell Rep Med.* 2021:2:100395. https://doi.org/10.1016/j.xcrm.2021.100395.

Dobin A *et al.* STAR: ultrafast universal RNA-Seq aligner. *Bioinformatics*. 2013:29:15–21. https://doi.org/10.1093/bioinformatics/bts635.

Eddy EM. Regulation of gene expression during spermatogenesis. *Semin Cell Dev Biol*. 1998:9:451–457. https://doi.org/10.1006/scdb.1998.0201.

Floriot S *et al.* CEP250 is required for maintaining centrosome cohesion in the germline and fertility in male mice. *Front Cell Dev Biol*. 2021:9:754054. https://doi.org/10.3389/fcell.2021.754054.

Fujihara M, Kim SM, Minami N, Yamada M, Imai H. Characterization and in vitro culture of male germ cells from developing bovine testis. *J Reprod Dev.* 2011:57:355–364. https://doi.org/10.1262/jrd.10-185M.

Gao X *et al.* Long read genome assemblies complemented by single cell RNA-Sequencing reveal genetic and cellular mechanisms underlying the adaptive evolution of yak. *Nat Commun*. 2022:13:4887. https://doi.org/10.1038/s41467-022-32164-9.

Germain PL, Sonrel A, Robinson MD. pipeComp, a general framework for the evaluation of computational pipelines, reveals performant single cell RNA-Seq preprocessing tools. *Genome Biol*. 2020:21:227. https://doi.org/10.1186/s13059-020-02136-7.

Gewaily MS *et al.* Comparative expression of cell adhesion molecule1 (CADM1) in the testes of experimental mice and some farm animals. *Acta Histochem.* 2020:122:151456. https://doi.org/10.1016/j.acthis.2019.151456.

Green CD *et al.* A comprehensive roadmap of murine spermatogenesis defined by single-cell RNA-Seq. *Dev Cell*. 2018:46:651–667.e10. https://doi.org/10.1016/j.devcel.2018.07.025.

Grey C *et al.* Mouse PRDM9 DNA-binding specificity determines sites of histone H3 lysine 4 trimethylation for initiation of meiotic recombination. *PLoS Biol*. 2011:9:e1001176. https://doi.org/10.1371/journal.pbio.1001176.

Griswold MD. 50 years of spermatogenesis: sertoli cells and their interactions with germ cells. *Biol Reprod*. 2018:99:87–100. https://doi.org/10.1093/biolre/ioy027.

Guo J *et al.* The adult human testis transcriptional cell atlas. *Cell Res*. 2018:28:1141–1157. https://doi.org/10.1038/s41422-018-0099-2.

Guo JT *et al.* Single-cell analysis of the developing human testis reveals somatic niche cell specification and fetal germline stem cell establishment. *Cell Stem Cell*. 2021:28:764–778.e4. https://doi.org/10.1016/j.stem.2020.12.004.

Gustavsson I, Switonski M, Larsson K, Ploen L. Synaptonemal Complex analysis of spermatocytes in hybrids of silver fox and blue fox. *J Hered*. 1988:79:338–343. https://doi.org/10.1093/oxfordjournals.jhered.a110525.

Heller CG, Clermont Y. Spermatogenesis in man - an estimate of its duration. *Science*. 1963:140:184–186. https://doi.org/10.1126/science.140.3563.184.

Jeffares DC *et al.* Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. *Nat Commun.* 2017:8:14061. https://doi.org/10.1038/ncomms14061.

Jiang T *et al.* Long-read-based human genomic structural variation detection with cuteSV. *Genome Biol*. 2020:21:189. https://doi.org/10.1186/s13059-020-02107-y.

Johnston SE. Understanding the genetic basis of variation in meiotic recombination: past, present, and future. *Mol Biol Evol*. 2024:41:msae112. https://doi.org/10.1093/molbev/msae112.

Kubota H, Avarbock MR, Brinster RL. Culture conditions and single growth factors affect fate determination of mouse spermatogonial stem cells. *Biol Reprod.* 2004:71:722–731. https://doi.org/10.1095/biolreprod.104.029207.

Larson EL, Keeble S, Vanderpool D, Dean MD, Good JM. The composite regulatory basis of the large X-effect in mouse speciation. *Mol Biol Evol.* 2016:34:282–295. https://doi.org/10.1093/molbev/msw243.

Lau X, Munusamy P, Ng MJ, Sangrithi M. Single-cell RNA sequencing of the cynomolgus macaque testis reveals conserved transcriptional profiles during mammalian spermatogenesis. *Dev Cell*. 2020:54:548–566.e7. https://doi.org/10.1016/j.devcel.2020.07.018.

Lemmon AR, Kirkpatrick M. Reinforcement and the genetics of hybrid incompatibilities. *Genetics*. 2006:173:1145–1155. https://doi.org/10.1534/genetics.105.048199.

Li YC, Wang GW, Xu SR, Zhang XN, Yang QE. The expression of histone methyltransferases and distribution of selected histone methylations in testes of yak and cattle-yak hybrid. *Theriogenology*. 2020:144:164–173. https://doi.org/10.1016/j.theriogenology.2020.01.001.

Lie PP, Cheng CY, Mruk DD. Coordinating cellular events during spermatogenesis: a biochemical model. *Trends Biochem Sci*. 2009:34:366–373. https://doi.org/10.1016/j.tibs.2009.03.005.

MacEachern S, McEwan J, Goddard M. Phylogenetic reconstruction and the identification of ancient polymorphism in the Bovini tribe (bovidae, bovinae). *Bmc Genomics*. 2009:10:177. https://doi.org/10.1186/1471-2164-10-177.

Mack KL, Nachman MW. Gene regulation and speciation. *Trends Genet*. 2017:33:68–80. https://doi.org/10.1016/j.tig.2016.11.003.

Magnusdottir E *et al.* A tripartite transcription factor network regulates primordial germ cell specification in mice. *Nat Cell Biol*. 2013:15:905–915. https://doi.org/10.1038/ncb2798.

Mihola O, Trachtulec Z, Vlcek C, Schimenti JC, Forejt J. A mouse speciation gene encodes a meiotic histone H3 methyltransferase. *Science*. 2009:323:373–375. https://doi.org/10.1126/science.1163601.

Mipam T *et al.* Single-cell transcriptome analysis and in vitro differentiation of testicular cells reveal novel insights into male sterility of the interspecific hybrid cattle-yak. *Bmc Genomics*. 2023:24:149. https://doi.org/10.1186/s12864-023-09251-2.

Moura AA, Souza CEA, Erickson BH. Early prepubertal testis criteria, seminiferous epithelium and hormone concentrations as related to testicular development in beef bulls. *Anim Reprod Sci.* 2011:124:39–47. https://doi.org/10.1016/j.anireprosci.2011.01.016.

Mu W, Starmer J, Fedoriw AM, Yee D, Magnuson T. Repression of the soma-specific transcriptome by polycomb-repressive complex 2 promotes male germ cell development. *Genes Dev*. 2014:28:20562069. https://doi.org/10.1101/gad.246124.114.

Murat F *et al.* The molecular evolution of spermatogenesis across mammals. *Nature*. 2023:613:308–316. https://doi.org/10.1038/s41586-022-05547-7.

Niayale R, Cui Y, Adzitey F. Male hybrid sterility in the cattle-yak and other bovines: a review. *Biol Reprod*. 2021:104:495–507. https://doi.org/10.1093/biolre/ioaa207.

Oatley MJ, Kaucher AV, Yang QE, Waqas MS, Oatley JM. Conditions for long-term culture of cattle undifferentiated spermatogonia. *Biol Reprod.* 2016:95:14–14. https://doi.org/10.1095/biolreprod.116.139832.

Orr HA. Dobzhansky, bateson, and the genetics of speciation. *Genetics*. 1996:144:1331–1335. https://doi.org/10.1093/genetics/144.4.1331.

Pathak S, Kieffer NM. Sterility in hybrid cattle. 1. Distribution of constitutive heterochromatin and nucleolus organizer regions in somatic and meiotic chromosomes. *Cytogenet Cell Genet.* 1979:24:42–52. https://doi.org/10.1159/000131355.

Powers NR *et al.* Sexual dimorphism in the meiotic requirement for PRDM9: a mammalian evolutionary safeguard. *Sci Adv.* 2020:6:eabb6606. https://doi.org/10.1126/sciadv.abb6606.

Qiu Q *et al.* The yak genome and adaptation to life at high altitude. *Nat Genet*. 2012:44:946–949. https://doi.org/10.1038/ng.2343.

Qiu X *et al.* Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. 2017:14:979–982. https://doi.org/10.1038/nmeth.4402.

Rosen BD *et al. De novo* assembly of the cattle reference genome with single-molecule sequencing. *Gigascience*. 2020:9:giaa021. https://doi.org/10.1093/gigascience/giaa021.

Royal MD *et al.* Declining fertility in dairy cattle: changes in traditional and endocrine parameters of fertility. *Anim Sci*. 2000:70:487–501. https://doi.org/10.1017/S1357729800051845.

Runemark A, Moore EC, Larson EL. Hybridization and gene expression: beyond differentially expressed genes. *Mol Ecol.* 2025:34:e17303. https://doi.org/10.1111/mec.17303.

Sedlazeck FJ *et al.* Accurate detection of complex structural variations using single-molecule sequencing. *Nat Methods*. 2018:15:461–468. https://doi.org/10.1038/s41592-018-0001-7.

Segura-Wang M, Onishi-Seebacher M, Stütz AM, Mardin BR, Korbel JO. Systematic identification of determinants for single-strand annealing-mediated deletion formation in *Saccharomyces cerevisiae*. *G3 (Bethesda)*. 2017:7:3269–3279. https://doi.org/10.1534/g3.117.300165.

Shah MA *et al.* Isolation and characterization of spermatogenic cells from cattle, yak and cattleyak. *Anim Reprod Sci.* 2018:193:182–190. https://doi.org/10.1016/j.anireprosci.2018.04.067.

Shami AN *et al.* Single-cell RNA sequencing of human, macaque, and mouse testes uncovers conserved and divergent features of mammalian spermatogenesis. *Dev Cell*. 2020:54:529–547.e12. https://doi.org/10.1016/j.devcel.2020.05.010.

Shen R *et al.* Genomic structural variation-mediated allelic suppression causes hybrid male sterility in rice. *Nat Commun*. 2017:8:1310. https://doi.org/10.1038/s41467-017-01400-y.

Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. *Genome Res*. 2017:27:491–499. https://doi.org/10.1101/gr.209601.116.

Sohni A *et al.* The neonatal and adult human testis defined at the single-cell level. *J Urol.* 2019:201:E157–E157. https://doi.org/10.1016/j.celrep.2019.01.045.

Song Y, Miao Z, Brazma A, Papatheodorou I. Benchmarking strategies for cross-species integration of single-cell RNA sequencing data. *Nat Commun*. 2023:14:6495. https://doi.org/10.1038/s41467-023-41855-w.

Soto DA, Ross PJ. Similarities between bovine and human germline development revealed by single-cell RNA sequencing. *Reproduction*. 2021:161:239–253. https://doi.org/10.1530/REP-20-0313.

Staub C, Johnson L. Review: spermatogenesis in the bull. *Animal*. 2018:12:s27–s35. https://doi.org/10.1017/S1751731118000435.

Stévant I *et al.* Deciphering cell lineage specification during male sex determination with single-cell RNA sequencing. *Cell Rep.* 2018:22:1589–1599. https://doi.org/10.1016/j.celrep.2018.01.043.

Stuart KC, Edwards RJ, Sherwin WB, Rollins LA. Contrasting patterns of single nucleotide polymorphisms and structural variation across multiple invasions. *Mol Biol Evol*. 2023:40:msad046. https://doi.org/10.1093/molbev/msad046.

Stuart T *et al.* Comprehensive integration of single-cell data. *Cell*. 2019:177:1888–1902.e21. https://doi.org/10.1016/j.cell.2019.05.031.

Takashima S, Shinohara T. Culture and transplantation of spermatogonial stem cells. *Stem Cell Res.* 2018:29:46–55. https://doi.org/10.1016/j.scr.2018.03.006.

Trapnell C *et al.* The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat Biotechnol*. 2014:32:381–386. https://doi.org/10.1038/nbt.2859.

Tumennasan K *et al.* Fertility investigations in the F-1 hybrid and backcross progeny of cattle (Bos Taurus) and yak (B-grunniens) in Mongolia. *Cytogenet Cell Genet.* 1997:78:69–73. https://doi.org/10.1159/000134633.

Turner JM. Meiotic silencing in mammals. *Annu Rev Genet*. 2015:49:395–412. https://doi.org/10.1146/annurev-genet-112414-055145.

Wang X *et al.* Single-cell RNA sequencing and UPHLC-MS/MS targeted metabolomics offer new insights into the etiological basis for male cattle-yak sterility. *Int J Biol Macromol*. 2023:253:126831. https://doi.org/10.1016/j.ijbiomac.2023.126831.

Wang Z *et al.* Sertoli cell-only phenotype and scRNA-Seq define PRAMEF12 as a factor essential for spermatogenesis in mice. *Nat Commun*. 2019:10:5196. https://doi.org/10.1038/s41467-019-13193-3.

Wrobel KH, Suss F. Identification and temporospatial distribution of bovine primordial germ cells prior to gonadal sexual differentiation. *Anat Embryol (Berl).* 1998:197:451–467. https://doi.org/10.1007/s004290050156.

Wu SX, Wan RD, Wang GW, Zhang YW, Yang QE. Comparative proteomic analysis identifies differentially expressed proteins associated with meiotic arrest in cattle-yak hybrids. *Proteomics*. 2023:23:e2300107. https://doi.org/10.1002/pmic.202300107.

Yang F *et al.* Refined genome-wide comparative map of the domestic horse, donkey and human based on cross-species chromosome painting: insight into the occasional fertility of mules. *Chromosome Res*. 2004:12:65–76. https://doi.org/10.1023/B:CHRO.0000009298.02689.8a.

You S *et al.* A toxin-antidote system contributes to interspecific reproductive isolation in rice. *Nat Commun*. 2023:14:7528. https://doi.org/10.1038/s41467-023-43015-6.

Zakharyevich K *et al.* Temporally and biochemically distinct activities of Exo1 during meiosis: double-strand break resection and resolution of double Holliday junctions. *Mol Cell*. 2010:40:1001–1015. https://doi.org/10.1016/j.molcel.2010.11.032.

Zhang H, Liu B, Qiu Y, Fan JF, Yu SJ. Pure cultures and characterization of yak sertoli cells. *Tissue Cell.* 2013:45:414–420. https://doi.org/10.1016/j.tice.2013.07.004.

Zhang S *et al.* Structural variants selected during yak domestication inferred from long-read whole-genome sequencing. *Mol Biol Evol.* 2021:38:3676–3680. https://doi.org/10.1093/molbev/msab134.

Zhang YW, Wu SX, Wang GW, Wan RD, Yang QE. Single-cell analysis identifies critical regulators of spermatogonial development and differentiation in cattle-yak bulls. *J Dairy Sci.* 2024:107:7317–7336. https://doi.org/10.3168/jds.2023-24442.

Zhao SS, Wu SX, Jia GX, Abulizi W, Yang QE. Localization and expression of SLX4 in the testis of sterile male cattle-yak. *Reprod Domest Anim*. 2023:58:679–687. https://doi.org/10.1111/rda.14338.

Zickler D, Kleckner N. Recombination, pairing, and Synapsis of homologs during meiosis. *Cold Spring Harb Perspect Biol*. 2015:7:a016626. https://doi.org/10.1101/cshperspect.a016626.